

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on : January 19, 2003, 19:50:40 ; search time 2218 Seconds

(without alignments)
9324.453 Million cell updates/sec

Title: OS-09-844-864-1

Perfect score: 1277

Sequence: 1 aaggccggcaggcgccggaa.....acaaaaaaaaaaaaaa 1277

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthun:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_mam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
C 1	646.2	50.6	673 12	BG071693
C 2	643	50.4	675 10	BB641267
C 3	642.4	50.3	666 10	BB264222
C 4	580.4	45.5	593 12	BG084538
C 5	570.4	44.7	594 12	BF471866
C 6	497.2	38.9	525 10	BB704019

ALIGNMENTS

Date not Applicable

RESULT 1
BG071693_c

DERIVED FROM H3102B03-3 NIA Mouse 15K cDNA Clone Set mRNA sequence.

ACCESSION BG071693

VERSION BG071693.1 GI:12554262

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus; 1 (bases 1 to 673)

REFERENCE Kargul, G.J., Dudekula, D.B., Qian, Y., Lim, M.K., Jaradat, S.A., Tanaka, T.S., Carter, M.G. and Ko, M.S.H.

AUTHORS Verification and initial annotation of NIA mouse 15K cDNA clone set

TITLE Unpublished (2001)

JOURNAL Other ESTs: H3102B03-5

COMMENT Contact: George J. Kargul

Laboratory of Genetics National Institute on Aging/National Institutes of Health 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA Email: cdna@lgsun.grc.nia.nih.gov

This clone set has been freely distributed to the community. Please visit <http://lgsun.grc.nia.nih.gov/cDNA/15k.html> for details.

Plate: H3102 row: B column: 03 Seq primer: -21M13 Forward High quality sequence stop: 673 POLYA=Yes.

FEATURES
source

Location/Qualifiers

1. . 673 /organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="niaEST:H3102B03-3"
/db_xref="taxon:10090"
/clone_lib="NIA Mouse 15K cDNA Clone Set"
/sex="Clones arrayed from a variety of cDNA libraries"
/dev_stage="Clones arrayed from a variety of cDNA libraries"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This clone is among a rearrrayed set of 15,247 clones from 11 embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos , and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, Proc. Natl. Acad. Sci. U.S.A., 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

BASE COUNT

137 a 191 c 173 g 172 t

ORIGIN

Query Match 50.6%; Score 646.2; DB 12; Length 673;
Best Local Similarity 99.3%; Pred. No. 7.1e-103;
Matches 670; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 1125 AACGTTCTGCTAGATGGGCTAATGGAATGGACAAGTGAGCTTCCTCCCTCTTCACC 1184
Db 134 AACGTCTGTGCTAGATGGGCTAATGGAATGGACAAGTGAGCTTCCTCCCTCTTCACC 75
QY 1185 TCTCCCTTCCAATCTTCATGACAGACAGTGTACTTGGATATAAGCCTGTGATA 1244
Db 74 TCTCCCTTCCAATCTTCATGACAGACAGTGTACTTGGATATAAGCCTGTGATA 16
QY 1245 AAAGGTATGCAAAC 1259
Db 15 AAAGGTATTGCAAAC 1

RESULT 2

BB641267 BB641267 RIKEN full-length enriched, 10 days neonate cortex Mus

DEFINITION BB641267 mRNA linear EST 26-OCT-2001

LOCUS BB641267 RIKEN full-length enriched, 10 days neonate cortex Mus

ACCESSION BB641267

VERSION BB641267.1 GI:16476392

KEYWORDS EST.

SOURCE

ORGANISM MUS musculus

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 675)

Akrawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.

JOURNAL RIKEN Mouse ESTs (Akrawa,T., et al. 2001)

COMMENT Unpublished (2001)
Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp/
URL: http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. . 10 (10), 1617-1630 (2000)
wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanuki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuzura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. . 10 (11), 1757-1771 (2000)
Komio,H., Furukoshi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y.

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

mouse tissues.

FEATURES
source

Location/Qualifiers

1. . 675 /organism="Mus musculus"

/db_xref="taxon:10090"
 /clone_lib="RIKEN full-length enriched, 10 days neonate cortex"
 /tissue_type="cortex"
 /dev_stage="10 days neonate"
 /lab_host="DHL0B"

/note="Site_1: Sall; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer 15, GAGAGAGAAGGATCCAGAGCTCTTTTTTTTVN 3', cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence 15', GAGAGAGATTCGAGTTAATTAAATTCCCCCCCCCCC 3'. cDNA was cleaved with XbaI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC 1."

BASE COUNT	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah	ai	aj	ak	al	am	an	ao	ap	aq	ar	as	at	au	av	aw	ax	ay	az	ba	ca	da	ea	fa	ga	ia	aa	ab	ac	ad	ae	af	ag	ah</th
------------	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	--------

/clone_lib="RIKEN full-length enriched, 10 days neonate cortex"
 /tissue_type="cortex"
 /dev_stage="10 days neonate"
 /lab_host="DH10B"
 /note="Site_1: Sall; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATCTCGAGTTAATTAAATCCCCCCCCCCC 3']. cDNA was cleaved with xhol and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."
 BASE COUNT 167 a 171 c 192 g 136 t
 ORIGIN
 Query Match 50.3%; Score 642.4; DB 10; Length 666;
 Best Local Similarity 99.6%; Pred. No. 3.2e-102;
 Matches 665; Conservative 0; Mismatches 1; Indels 2; Gaps 2;
 QY 588 GCAGGGCGATGTTCAAGCTGCAGGGCAGGCCGGGGGGAGGAGCACCACGGAGGA 647
 Db 1 GGAGGGCGATGTTCAAGCTGCAGGGCAGGCCGGGGGGAGGAGCAGCCACCGGAGGA 60
 QY 648 CCGGACACAGTGTGGCGCGATGCAGTCAGGCTGGAGCGAGGCCATGTCCTGCAGC 707
 Db 61 CCGGACACAGTGTGGCGCGATGCAGTCAGGCTGGAGCGAGGCCATGTCCTGCAGC 120
 QY 708 AGAGATGGCTCAGGACCCGGTAGCTGGATGCCCTCGAGAACCGAGGCCCTCCCGCAAAG 767
 Db 121 AGAGATGGCTCAGGACCCGGTAGCTGGATGCCCTCGAGAACCGAGGCCCTCCCGCAAAG 180
 QY 768 CACGGAGCAGGACAAAGGGAGGGCTGGTTCCAGTCTAGAGCAGAACGGCTACTA 827
 Db 181 CACGGAGCAGGACAAAGGGAGGGCTGGTTCCAGTCTAGAGCAGAACGGCTACTA 240
 QY 828 TCACTGCAAGGACTGCAAATCCGGTGGAGAGGCCCTATGTGGTGTGAGGGAC 887
 Db 241 TCACTGCAAGGACTGCAAATCCGGTGGAGAGGCCCTATGTGGTGTGAGGGAC 300
 QY 888 CAGTAAGGTGTTACTTCAAACAGTCTGCCGAGGTGTGAGAAATCCTACAAACCTTACA 947
 Db 301 CAGTAAGGTG-TACTTCAAACAGTCTGCCGAGGTGTGAGAAATCCTACAAACCTTACA 359
 QY 948 GAGTGGAGGACATCACCTGCAAAAGTTGTAAGAAACTAGATGTGGAGATGCAAGGACA 1007
 Db 360 GAGTGGAGGACATCACCTGCAAGTGTGAGAAACTAGATGTGGAGATGCAAGGACA 419
 QY 1008 TTCGCCACGGACCCATAACGCCCTCGGACAGCCCTAACGCCCTACGGCAAGGACA 1067
 Db 420 AACGCCACGGACAGCACCTCAGCTCAAATACATCATTTAGTGAGAGTCGAAAG 479
 QY 1128 CGTTTCTGCTAGATGGGAATGGACAAGTGAGCTTCAGCTCAAATACATCATTTAGTGAGAGTCGAAAG 1127
 Db 540 CGTTTCTGCTAGATGGGCTAATGGAATGGACAAGTGAGCTTCAGCTCAAATACATCATTTAGTGAGAGTCGAAAG 599
 QY 1188 TCCCTTCCAATTCTCATGACAGACAGACTGTGTTACTTGAGATAAGCCGTGAAATAAAGCCCTGTGAAATAA 1247
 Db 600 TCCCTTCCAATTCTCATGACAGACAGACTGTGTTACTTGAGATAAGCCCTGTGAAATAAAGCCCTGTGAAATAA 658

QY 1248 GGTATIGC 1255
 Db 659 GGTATTGC 666

RESULT 4
 BG084538 mRNA 593 bp linear EST 26-JAN-2001
 LOCUS H3102B03-5 NIA Mouse 15K cDNA clone Set Mus musculus CDNA clone
 DEFINITION H3102B03 5', mRNA sequence.
 ACCESSION BG084538
 VERSION BG084538.1 GI:12567102
 KEYWORDS EST.
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 593)
 AUTHORS Kargul, G.J., Dudekula, D.B., Qian, Y., Lim, M.K., Jaradat, S.A., Tanaka, T.S., Carter, M.G. and Ko, M.S.H.
 TITLE Verification and initial annotation of NIA mouse 15K cDNA clone set
 JOURNAL Unpublished (2001)
 COMMENT Other refs. H3102B03-3
 Contact: George J. Kargul
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 400, Baltimore, MD 21224-6820, USA
 Email: cdna@igsun.grc.nia.nih.gov
 This clone set has been freely distributed to the community. Please visit <http://igsun.grc.nia.nih.gov/cDNA/15k.html> for details.
 Plate: H3102 row: B column: 03
 Seq primer: -21M13 Reverse
 High quality sequence stop: 593
 POLY(A)
 FEATURES Location/Qualifiers
 source 1. .593
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="nlaEST:H3102B03-5"
 /db_xref="taxon:10090"
 /clone="H3102B03"
 /clone="NIA Mouse 15K cDNA Clone Set"
 /sex="Clones arrayed from a variety of cDNA libraries"
 /dev_stage="Clones arrayed from a variety of cDNA libraries"
 /lab_host="DH10B"
 /note="vector: pSPORT1; Site_1: Sall; Site_2: NotI; This clone is among a rearraigned set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb: All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex development, 127: 1967-1978." Mol Genet 7: 1967-1978."
 BASE COUNT 142 a 165 c 192 g 94 t
 ORIGIN
 Query Match 45.5%; Score 580.4; DB 12; Length 593;
 Best Local Similarity 99.7%; Pred. No. 2e-91;
 Matches 592; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 512 CGGCATCCTCGGGACCCGGAGGCCAGGGTGGCCGGAAAGCGGTCC 571

Db	1	CGGCATCTCGGGACCCGGAACCGGAGCCGAGAGGTGGCCGGAGGAAGCGGCC	source	1. .594
Qy	572	CCCAGCCGGAAGCAGGAGGGGGATTTTCAGGCTGCAGGGCAGGCCGGAGC	/organism="Mus musculus"	/strain="C57BL/6J"
Db	61	CCGCGCGGAAGCAGGAGGGGGATTCAGGCTGCAGGGCAGGCCGGAGCAGC	/db_xref="taxon:10090"	/clone="UI-M-BH3-awu-b-08-0-UI"
Qy	632	AGCCACCACCGGAGGACCGGAAACTGAGATGGCTCAGGACCGGGAGGAGC	/clone_lib="NIH_BMAP_M_S4"	/dev_stage="27-32 days"
Db	121	AGCCACCACCGGAGGACCGGAAACTGAGATGGCTCAGGACCGGGAGGAGC	/lab_host="DH10B (Life Technologies)"	/note="vector: pr7T3D-Pac (Pharmacia) with a modified
Qy	692	AGCCATGTCCTGCCAGAGATGGCTCAGGACCGGGATCGGATGCCCTCGA	/polylinker; Site_1: Not I; Site_2: Eco RI; The	NIH_BMAP_M_S4 library is a subtracted library of a series,
Db	181	GCCCATGTCCTGCCAGAGATGGCTCAGGACCGGGATCGGATGCCCTCGA	ultimately derived from a mixture of individually tagged	normalized libraries from ten regions of the mouse brain
Qy	752	AGGCTCCCCGAAAGCACGGAGCAGGACAAGGGCCCTGGTTCCAGTCTTA	(cerebellum, brain stems, olfactory bulbs, hypothalamus,	cortex, amygdala, basal ganglia, pineal gland, striatum,
Db	241	AGGCTCCCCGAAAGCACGGAGCAGGACAAGGGCCCTGGTTCCAGTCTTAG	hippocampus) after a series of subtractions to reduce the	representation of cDNAs from which ESTs had already been
Qy	812	AGAACTACGGCTACTATCACTGCAAAGGACTGCAAATCCGGGGATGTT	generated. The following serially subtracted libraries	were generated in this process: NIH_BMAP_M_S4,
Db	301	AGAAGTACGGCTACTATCACTGCAAAGGACTGCAAATCCGGGGATGTT	NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S3.1,	NIH_BMAP_M_S2, NIH_BMAP_M_S1. The subtracted library
Qy	872	CGTGTGTCAGGGCACCAAGTAAGGTGTTACTTCAAACAGTTCTGCCAG	(NIH_BMAP_M_S4) was constructed as follows: PCRamplified	CDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and
Db	361	GTTGTCAGGGCACCAAGTAAGGTGTTACTTCAAACAGTTCTGCCAG	NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived	was used as a driver in a hybridization with a pool of
Qy	932	TCTACAAACCTTACAGAGTGGGGACATCACCTGTCAAAAGTTGTA	the NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1	libraries in the form of single-stranded circles. The
Db	420	TCCATACAAACCTTACAGAGTGGGGACATCACCTGTCAAAAGTTG	remaining single-stranded circles (subtracted library)	was purified by hydroxyapatite column chromatography,
Qy	992	GCCTGCCAGTCAGTTGCCAACGTTGACCTTAACGCCCCATGGCAAG	converted to double-stranded circles and electroporated	into DH10B bacteria (Lifetechnologies) to generate the
Db	480	GCCTGCCAGTCAGTTGCCAACGTTGACCTTAACGCCCCATGGCAAG	described (Bonaldo, Lennon and Soares, Genome Research	NIH_BMAP_M_S4 library. This procedure has been previously
Qy	1052	GGGAGATGCAAGGACAACGCCCTGTCCTGGCACACCTTCAGCTCAAA	6:791-806, 1996)	described (Bonaldo, Lennon and Soares, Genome Research
Db	540	GGGAGATGCAAGGACAACGCCCTGTCCTGGCACACCTTCAGCTCAAA	593	593, 1996)
RESULT	5	Query Match	44.78; Score 570.4; DB 12; Length 594;	
BF471866/c		Best Local Similarity	99.5%; Pred. No. 1.1e-89;	
LOCUS	BF471866	Matches	593; Conservative 0; Mismatches 1; Indels 2; Gaps 2;	
DEFINITION	UT-M-BH3-awu-b-08-0-UI.r1	Matches	593; Conservative 0; Mismatches 1; Indels 2; Gaps 2;	
ACCESSION	NIH_BMAP_M_S4	Definition	NIH_BMAP_M_S4 Mus musculus cDNA clone	
VERSION	BF471866	Version	5	
KEYWORDS	EST.	Origin	NIH_BMAP_M_S4	
SOURCE	EST.	BASE COUNT	105 a 184 c 158 g 147 t	
ORGANISM	house mouse.	ORIGIN		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 594)			
AUTHORS	Bonaldo, M.F., Lennon, G. and Soares, M.B.			
TITLE	Normalizat <u>ion</u> and subtraction: two approaches to facilitate gene discovery			
JOURNAL	Genome Res. 6 (9), 791-806. (1996)			
ISSNLINE	0270-4472			
COMMENT	Contact: Chin, H			
	National Institute of Mental Health			
	6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD			
	20892-9643, USA			
	Tel: 301 443 1706			
	Fax: 301 443 9890			
	Email: MEST@mail.nih.gov			
CDNA Library Preparation:	M.B. Soares Lab Clone distribution:			
Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It	should be noted that Bento Soares is generating a small number of			
additional specialized non-redundant arrays of BMAP cDNAs whose	availability will be considered under appropriate and limited			
collaborative arrangements				
Seq primer: M13 Reverse.				
FEATURES	Location/Qualifiers			

	RESULT 6	BB704019	525 bp mRNA linear EST 11-OCT-2001
LOCUS	BB704019	RIKEN full-length enriched, in vitro fertilized eggs Mus	
DEFINITION	BB704019	musculus cDNA clone 7420459B08 3', mRNA sequence.	
ACCESSION	BB704019	BB704019.1 GI:16052854	
VERSION	BB704019.1	EST.	
KEYWORDS			
SOURCE		house mouse.	
ORGANISM	Mus musculus	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE	1	(bases 1 to 525)	
AUTHORS	Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Watanuki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.		
TITLE	RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.)		
JOURNAL	Unpublished (2001)		
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.go.jp/ URL: http://genome.gsc.riken.go.jp/		
BASE COUNT	142 a 134 c 129 g 120 t		
ORIGIN			
Query Match	38.9%	Score 497.2;	DB 10;
Best Local Similarity	99.0%	Pred. No. 6.1e-77;	Length 525;
Matches	521	Mismatches 3;	Indels 2;
Conservative	0	Gaps 2;	
Db	732	TTCGGATGCCCTCGAGACCAGCAGGCTCCCCAAGCAGGAGCACAGGGCCT	791
Qy	1	TTCCGGATGCCCTCGAGACCAGGCTCCCCAAGCAGGAGCACAGGCCT	60
Db	792	GGGTTCAGTCTTAGAGCAGAAATCGCTACTATCACTGCAAGGACTGCAAATCCG	851
Qy	61	GGCTTTCAGTCTTAGAGCAGAAATCGCTACTATCACTGCAAGGACTGCAAATCCG	120
Db	852	GTGGGAGGGCCTATGTGTGGTGTGCAAGGGCACCAGTAAGGTACTTCAAACAGT	911
Qy	121	GTGGGAGGGCCTATGTGTGGTGTGCAAGGGCACCAGTAAGGT-TACTTCAAACAGT	179
Db	912	TCTGCCGAGTGTGAGAAATCCCTACACCCCTACAGAGTGGAGGACATCACCTGTCAA	971
Qy	180	TCTGCCGAGTGTGAGAAATCCCTACACCCCTACAGAGTGGAGGACATCACCTGTCAA	239
Db	972	GTTGTAAGAAACTAGATGTGCCAGTCAGATTGCCACGGACCTAACGCC	1031
Qy	240	GTGTTAAAGAACTAGATGTGTGGAGATGCAAGGACAACGCCCTAACGCC	299
Db	1032	CCCATCGGCAAGACTGTGTGGAGATGCAAGGACAACGCCCTAACGCC	1091
Qy	300	CCCATCGGCAAGACTGTGTGGAGATGCAAGGACAACGCCCTAACGCC	359
Db	1092	TCAGCTCAAATACTACATCATTAGTGAGAGTCGAAAGGACAACGCCCTAACGCC	1151
Qy	360	TCAGCTCAAATACTACATCATTAGTGAGAGTCGAAAGGACAACGCCCTAACGCC	419
Db	420	GAATGGACAAAGTGTGGCTTCTCCCTCTCACCTCTCCCTTCACCTCTCC	479
Qy	1152	GAATGGACAAAGTGTGGCTTCTCCCTCTCACCTCTCCCTTCACCTCTCC	1211
Db	420	GAATGGACAAAGTGTGGCTTCTCCCTCTCACCTCTCCCTTCACCTCTCC	524
Db	480	GACAGTGTACTTGGATATAAGCCGTGTGATAAAAGGTATTGCAA	524
FEATURES	e mouse tissues.		
source	Location/Qualifiers		
	1. .525		
	/organism="Mus musculus"		
	/strain="C57BL/6J"		
	/db_xref="taxon:10090"		
	/clone="7420459B08"		
	/clone_lib="RIKEN full-length enriched, in vitro fertilized eggs"		
	/sex="female"		
	/tissue_type="in vitro fertilized eggs"		
	/dev_stage="egg"		
RESULT 7			
LOCUS	A1854700	499 bp mRNA linear EST 15-JUL-1999	
DEFINITION	UT-M-BH0-akc-d-12-0-UT.s1_NTH_BMAP_M_S1	Mus musculus cDNA clone	
ACCESSION	A1854700	UT-M-BH0-akc-d-12-0-UT.s1_NTH_BMAP_M_S1	
VERSION	A1854700.1	GI:5498606	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	

/lab_host="DH10B"
 /note="Site_1: Sali; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGAGATTCCTCGAGTTAAATCCCCCCCC 3']. cDNA was cloned into the XbaI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: Sali; 3' end: BamHI"

REFERENCE	1	(bases 1 to 499)
AUTHORS	Bonaldo, M.F., Lennon, C. and Soares, M.B.	
TITLE	Normalization and subtraction: two approaches to facilitate gene discovery	
JOURNAL	Genome Res. 6 (9), 791-806 (1996)	
MEDLINE	97044477	
COMMENT	Contact: Chin, H National Institute of Mental Health 6001 Executive Blvd., Room 7N-7190, MSC 9643, Bethesda, MD 20892-9643, USA Tel: 301 443 1706 Fax: 301 443 9890 Email: mEST@mail.nih.gov	
The sequence contained an oligo-dT track that was present in the normalized basal ganglia library preparation: M.B. Soares Lab Clone distribution: NIH BMAP cDNA clones will be made available by the means that is soon to be determined. When NIH determines the means for distribution of the BMAP cDNA clones, this record will be updated accordingly when that means is determined.		
Seq primer: M13 Forward POLYA=Yes		
FEATURES	Location/Qualifiers	
source	1. . 499 'organism="Mus musculus" 'strain="C57BL/6J" 'db_xref="taxon:10090" 'clone="UI-M-BHO-akc-d-12-0-UI" 'clone_1ib="NIH_BMAP_M_S1" 'dev_stage="27-32 days" 'lab_host="DH10B (Life Technologies)" 'note="vector: pRT3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; The NIH_BMAP_M_S1 library is a subtracted library derived from a mixture of normalized libraries from ten regions of the mouse brain (cerebellum, brain stems, olfactory bulbs, hypothalamus, cortex, amygdala, basal ganglia, pineal gland, striatum, hippocampus). The driver used for subtraction consisted of a pool of 20,000 cDNA clones obtained from non-normalized and normalized libraries of these ten regions of the mouse brain." TAG_LIB=NIH_BMAP_M_S1 TAG_TISSUE=basal-ganglia TAG_SEQ=TGTAC"	
BASE COUNT	115 a 120 c 117 g 147 t	
ORIGIN		
Query Match	37.1%	Score 473.8; DB 9; Length 499;
Best Local Similarity	99.2%	Pred. No. 7.2e-73;
Matches	497	Conservative 0; Mismatches 2; Indels 2; Gaps 2;
Qy	772	GAGCAGACAAGGAGCCCTGCCTTCAGTTCAAGCAGAAGTACGGCTACTATCAC 831
Db	499	GAGCAGACAAGGAGCCCTGCCTTCAGTTCAAGCAGAAGTACGGCTACTATCAC 440
Qy	832	TGCCAAGACTGCAAATCCGGTGGAGAGCCCTATGTGGTGTGCAGGGCACCAGT 891
Db	439	TGCCAAGACTGCAAATCCGGTGGAGAGCCCTATGTGGTGTGCAGGGCACCAGT 380
Qy	892	AAGCTGTACTTCAAACAGTTCTGCCAGTGTGAGAACCTAACAGT 951
Db	379	AAGGTG-TACTTCAAACAGTTCTGCCAGTGTGAGAACCTAACAGT 321
Qy	952	GGAGGACATCCACCTGTCAAAGTGTAAAGAACTAGATGTGCCAGTCAGATTG 1011
Db	320	GGAGGACATCCACCTGTCAAAGTGTAAAGATAGATGTGCCAGTCAGACTTCG 261
Qy	1012	CCACCTGGACCCCTAACGCCCATCGGCAAGACTGTGTGGAGATGCAAGGACAACG 1071
Db	260	CCACCTGGACCCCTAACGCCCATCGGCAAGACTGTGTGGAGATGCAAGGACAACG 201
FEATURES	Location/Qualifiers	
source	1. . 499	
REFERENCE	1072 CCTGTCCTGCCAGACCTTCAGCTCAAATCATCATTAGTGAGACTGAAACGTT 1131 Db 200 CCTGTCCTGCCAGACCTTCAGCTCAAATCATCATTAGTGAGACTGAAACGTT 141 Qy 1132 TCTGCTAGATGGGCTAATGGAATGGCAAGTGTGAGCTTCCCTCTCACCTCTCCC 1191 Db 140 TCTGCTAGATGGGCTAATGGAATGGCAAGTGTGAGCTTCCCTCTCACCTCTCCC 81 Qy 1192 TTGCAAATTCTCATGACAGACAGTGTACTTGATAAAGCCTGTGAATAAAGGTA 1251 Db 80 TTGCAAATTCTCATGACAGACAGTGTACTTGATAAAGCCTGTGAATAAAGGTA 22 Qy 1252 TTGCAAACAAAAAAA 1272 Db 21 TTGCAAACAAAAAAA 1	
RESULT	8	
BB703259	BB703259	
DEFINITION	RTKEN full-length enriched, in vitro fertilized eggs Mus mRNA sequence.	
ACCESSION	BB703259	
VERSION	BB703259.1	
KEYWORDS	GI:16052094	
SOURCE	EST.	
ORGANISM	Mus musculus	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 491)	
AUTHORS	Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Watanuki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.	
TITLE	RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al. 2001)	
JOURNAL	Unpublished (2001)	
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.go.jp URL: http://genome.gsc.riken.go.jp/	
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. . 10 (10), 1617-1630 (2000)	
wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Kira,A. and Hayashizaki,Y.	Riken integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. . 10 (11), 1757-1771 (2000)	
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.	Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)	
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.	e mouse tissues.	

	DEFINITION	BB704648 RIKEN full-length enriched, in vitro fertilized eggs Mus musculus CDNA clone 7420466107 3', mRNA sequence.
	ACCESSION	BB704648
	VERSION	BB704648.1 GI:16053483
	KEYWORDS	EST.
SOURCE	house mouse	
ORGANISM	Mus musculus	
JOURNAL	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.	
COMMENT		
TITLE		
BASE COUNT		
ORIGIN		
Query Match	35.1%; Score 448.4; DB 10; Length 491;	
Best Local Similarity	98.2%; Pred. No. 1.9e-68;	
Matches	485; Conservative 0; Mismatches 6; Indels 3; Gaps 3;	
QY	763 CAAAGCACGGAGCAGAACAGGAGCGCTGGCTTCCAGTTCTTAGAGCAGAAGTACGGC 822	
Db	1 CAAAGCACGGAGCAGAACAGGAGGCCCTGCGTTTCAGTTATAGCAGAAGTACGGC 60	
QY	823 TACTATCACTGCAAGACTGCAAATCCGGGGAGAGGCCTATGTGTGGTGTGAG 882	
Db	61 TACTATCACTGCAA-GACTGCAAATCCGGGGAGAGGCCTATGTGTGGTGTGAG 119	
QY	883 GGCACCACTGAAGGTGTACTCAACAGTTCTGCCGAGTGTGAGAAATCCTACACCC 942	
Db	120 GGCACCACTGAAGGTG-TACTTAAACAGTTCTGCCGAGTGTGAGAAATCCTACACCC 178	
QY	943 TTACAGAGTGGAGGACATCACCTGTCAAAGTTGTAAGAAACTAGATGTGCCCTGCCCAGT 1002	
Db	179 TTACAGAGTGGAGGACATCACCTGTCAAAGTTGTAAGAAACTAGATGTGCCCTGCCCAGT 238	
QY	1003 CAGATTGCCAACGTGGACCTAACGCCCATGGCAAGACTGTGTGGAGATGCAA 1062	
Db	239 CAGACTTCGCCAACGTGGACCTAACGCCCATGGCAAGACTGTGTGGAGATGCAA 298	
FEATURES	Location/Qualifiers	
source	1. .491	
	/organism="Mus musculus"	
	/strain="C57BL/6J"	
	/db_xref="taxon:10090"	
	/clone="7420466107"	
	/clone_lib="RIKEN full-length enriched, in vitro fertilized eggs"	
	/sex="female"	
	/tissue_type="in vitro fertilized eggs"	
	/dev_stage="egg"	
	/lab_host="DH10B"	
	/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAAGGATCCAAGAGCTCTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGAATTCCTCGAGTTATTAAATTACCCCCCCCCCCC 3']. cDNA was cloned into the XbaI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI"	
RESULT	9	
LOCUS	BB704648 BB704648	

			primer adapter of sequence [5' GAGAGAGATCTCGACTTAATTAAATCCCCCCCCCC 3']. cDNA was cloned into the XbaI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI"
BASE COUNT	144 a	121 c	107 g
ORIGIN	119 t		
Query Match	32.4%	Score 413.8;	DB 10;
Best Local Similarity	93.0%	Pred. No. 1.9e-62;	Length 491;
Matches	455;	Conservative	0;
		Mismatches	32;
		Indels	2;
		Gaps	2;
Oy	768	CACGGAGCAGGACAAGGAGCTCCACTTCTGGAGACTACGGCTACTA	827
Db	1	CACGGAGCAGGACAAGGAGCTCCACTTCTGGAGACTACGGCTACTA	60
Oy	828	TCACTGCAAGGACTGCCAAATCCGGGGAGAGCCCTATGTGGTGTGCAGGCAC	887
Db	61	TCA-AGCAAGGACTGCCAAATCCGGCTGGGAAACGCCTATGTCTGGAGTGTGCAGGCAC	119
QY	888	CAGTAAAGGTGTTACTTCACAAACAGTTCTGCCAGGTGTGAGAAATCCTACAAACCTTACA	947
Db	120	CAGTAAGATGTTACTTCACACTTCTGCCAGGTGTGAGAAATCCTACAAACCTTACA	179
Oy	948	GAGTGGAGGACATCACCTGTCAAAGTTGTAAGAAGAACATAGATGTGCCAGTCAGAT	1007
Db	180	GAATGGAGGACATCACCTCTCAAATGTAAAGAACATAAATGTGCTTCCCAGTCAGAC	239
Oy	1008	TTGCCACGTTGACCCCTAACGCCCATGGCAAGACTGTGTGGAGATGCAAGGACA	1067
Db	240	TTGCCACGTTGACCCCTAACGCCCATGGCAAGACTGTGTGGAGATGCAAGGACA	299
Oy	1068	AACGCCCTGCTGGACACCTAACGCCCATGGCAAGACTTCAGCTCAAAATACATCATTAGTGAGAGTCGAA	1127
Db	300	AACGCCCTGCTGGACACCTTCAGCTCAAAATACATCATTAGTGAGAGTCGAA	359
Oy	1128	CGTTTCTGCTAGATGGGCTAATGGAATGACAAGTGTGAGCTTCTCCCTCTCACCTCT	1187
Db	360	CGTTTCTGCTAGATGGGCTAATGGAATGACAAGTGTGAGCTTCTCCCTCTCACCTCT	419
Oy	1188	TCCCTTCCAATTCTCATGACAGACAGTGTACTGGATAAAGCTGTGAATAAAA	1247
Db	420	TCCCTTCCAATTCTCATGACAGACAGTGTACTGGATAAAGCTGTGAATAAAA	478
Oy	1248	GGTATTGCA 1256	
Db	479	GGTATTGTA 487	
RESULT	10		
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE	AU044294	436 bp mRNA linear EST 09-DEC-1998	
DEFINITION	AU044294	Mouse sixteen-cell-embryo CDNA Mus musculus cDNA clone J0917G09 3', mRNA sequence.	
ACCESSION	AU044294		
VERSION	AU044294.1	GI:3979844	
KEYWORDS	EST.		
SOURCE			
house mouse.			
Mus musculus			
ERATO/Bio Project, ERATO Japan Science and Technology Corporation (JST) WBG Marine East 12F, 2-6 Nakase, Mihama-ku, Chiba 261-71, Japan			
COMMENT	Contact: Hirofumi Doi Email: hd@bio.jst.go.jp.		
JOURNAL			
Doi Bioassymetry Project, ERATO Japan Science and Technology Corporation (JST) WBG Marine East 12F, 2-6 Nakase, Mihama-ku, Chiba 261-71, Japan			
FEATURES	Location/Qualifiers		
source	1. -436		
/organism="Mus musculus"			
/strain="C57BL/6J"			
/db_xref="taxon:10090"			
/clone="J0917G09"			
/clone_1ib="Mouse sixteen-cell-embryo cDNA /dev_stage="sixteen-cell-embryo"			
/note="Organ: mammary gland; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TTTTACCAATCTGAAGTGGGGAGCCGGAATGTTTTTTTTTTTTTTT adaptor (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Minoru Ko, Wayne State Univ. Library constructed and normalized by Bento Soares and M.Fatima Bonaldo."			
BASE COUNT	106 a	104 c	105 g
ORIGIN	121 t		
Query Match	31.8%	Score 405.8;	DB 9;
Best Local Similarity	97.0%	Pred. No. 5e-61;	Length 436;
Matches	424;	Conservative	0;
		Mismatches	12;
		Indels	1;
		Gaps	1;
Oy	823	TACTATCACTGCAAGGACTGCAAATCCGGTGGAGAGCCCTATGTGTGGTGTGCAG	882
Db	436	TATTACTGCAAGGACTGCAAATCCGGTGGAGAGCCCTATGTGTGGTGTGCAG	377
Oy	883	GGCACCGAGTAAGGTGTTACTTCAAACAGTTCTGCCAGTGTGAGAAATCCTACAA	942
Db	376	GGCCACCACTGAGTGTACTTCACACAGTCTGCCAGTGTGAGAAATCCTACAA	317
Oy	943	TTACAGACTGGAGACATCACCTGTCAAAGTGTAAAGAACTAGATGTGCCCTGCCAGT	1002
Db	316	TTACAGACTGGAGACATCACCTGTCAAAGTGTAAAGAACTAGATGTGCCCTGCCAGT	257
Oy	1003	CAGATTGCCACGTTGACCCCTAACGCCCATGGCAAGACTGTGTGGAGATGCAA	1062
Db	256	CAGACTTCGCCACGTTGACCCCTAACGCCCATGGCAAGACTGTGTGGAGATGCAA	197
Oy	1063	GGACAAACGCCCTGCTGGACAGCACCTTCAGCTCAAATACATCATTAGTGAGAGTC	1122
Db	196	GGACAAACGCCCTGCTGGACAGCACCTTCAGCTCAAATACATCATTAGTGAGAGTC	137
Oy	1123	GAAACGCCCTGCTAGATGGGCTAATGGAATGACAAGTGTGTGGAGCTTCTCCCTCTCAA	1182
Db	136	GAAACGCCCTGCTAGATGGGCTAATGGAATGACAAGTGTGTGGAGCTTCTCCCTCTCAA	77
Oy	1183	CCTCTCCCTTCCAATTCTCATGACAGACAGTGTACTGGATAAAGCTGTGAA	1242
Db	76	CCTCTCCCTTCCAATTCTCATGACAGACAGTGTACTGGATAAAGCTGTGAA	18
KEYWORDS			
SOURCE			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE	1 (bases 1 to 425)	425 bp mRNA linear EST 11-OCT-2001	
DEFINITION	BB705931	RIKEN full-length enriched, in vitro fertilized eggs Mus musculus cDNA clone 7420484Fl6 3', mRNA sequence.	
ACCESSION	BB705931		
VERSION	BB705931.1	GI:16054766	
KEYWORDS	EST.		
SOURCE			
ORGANISM			
Mus musculus			
ERATO/Bio Project, ERATO Japan Science and Technology Corporation (JST) WBG Marine East 12F, 2-6 Nakase, Mihama-ku, Chiba 261-71, Japan			
COMMENT	Contact: Hirofumi Doi Email: hd@bio.jst.go.jp.		
JOURNAL			
Doi Bioassymetry Project, ERATO Japan Science and Technology Corporation (JST) WBG Marine East 12F, 2-6 Nakase, Mihama-ku, Chiba 261-71, Japan			
FEATURES	Location/Qualifiers		

BASE COUNT		ORIGIN		Query Match		Score 31.0%; DB 9; Length 590;	
Best Local Similarity		Pred. No.		85.1%; 0; Mismatches		1.9e-59; 71; Indels 4; Gaps 3;	
Matches 464; Conservative							
QY	713	TGGCTCAGGACCCGGTGTGGATGCCCTCGAGACAGGCCCTCCCCCAAAGCACGG	772				
Db	542	TGGNCNTNAGACCCCCGGTTATTCGGATGCCCTGGAAACAGGCCNTCCCNCAANNCGGACC	483				
QY	773	AGCAGGACAAGGAGCGCCTGCCTTCCAGTCTTAGAGCAGAAGTACGG--CTACTATCA	830				
Db	482	AGNACAAAGGAGGGCTGCNTTTCCAGCTACGNCTACTATCA	423				
QY	831	CTGCAAGGACTGCAAATCCCGNTGGAGANCNCCTATGTGTGAGAACCTACACCTTACAGA	949				
Db	422	CTGCAAGGACTGCAAATCCCGNTGGAGANCNCCTATGTGTGAGAACCTACACCTTACAGA	363				
QY	890	GTAAGGTGTACTTCACACAGTCTGCCAGTGTGAGAACATAGATGTGCNTGCCAGTCAGCT	1009				
Db	362	ACTAAGGTGTACTTCACACAGTCTGCCAGTGTGAGAACATAGATGTGCNTGCCAGTCAGCT	303				
QY	950	GTGGAGGACATCACCTGTCAAACTGTAAGAAACTAGATGTGCCAGTCAGATT	1009				
Db	302	GTGGAGGACATCACCTGTCAAACTGTAAGAAACTAGATGTGCCAGTCAGATT	243				
QY	1010	CGCCCACGTGGACCCTAACGCCCCATCGGCAAGACTTGTTGGAGATGCAAGGACAA	1069				
Db	242	CGGCCACGTGGACCCTAACGCCCCATCGGCAAGACTTGTTGGAGATGCAAGGACAA	183				
QY	1070	CGCCCTGTCTGGACAGCACCTTCAGCTCAATACATCATTAGTAGTGAGGTGCAAACG	1129				
Db	182	CGCCTGTCTGGACAGCACCTTNAGCTCAATACATCATTAGTGAGGTGCAAACG	123				
QY	1130	TTTCTGTAGATGGGCTATGGAACTGAGCTTCTCCCTCTTCACCTCTC	1189				
Db	122	TTTCTGTAGATGGGCTATGGAACTGAGCTTCTCCCTCTTCACCTCTC	63				
QY	1190	CCTTTCCTAACATTCTCATGACAGACAGTGTACTGGATAAAGCTGTGAATAAAG	1248				
Db	62	CCTTTCCTAACATTCTCATGACAGACAGTGTACTGGATAAAGCTGTGAATAAAG	5				
RESULT 13		BASE COUNT		ORIGIN		Query Match	
BE946858						30.8%; DB 12; Length 450;	
LOCUS	BE946858					Best Local Similarity 99.7%; Pred. No. 8.2e-59;	
DEFINITION		450 bp mRNA	linear EST 03-OCT-2000			Matches 393; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
ACCESSION	UI-M-BH3-awu-b-08-0-01.s1	NIH_BMAP_M_S4	Mus musculus cDNA clone				
VERSION	BE946858	1	UI-M-BH3-awu-b-08-0-01.s1				
KEYWORDS	BE946858.1	GT:10524617					
EST							
SOURCE							
ORGANISM							
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.							
REFERENCE	1	(bases 1 to 450)	Bonaldo, M.F., Lennon, G. and Soares, M.B.				
AUTHORS			Normalization and subtraction: two approaches to facilitate gene discovery				
JOURNAL			Genome Res. 6 (9), 791-806 (1996)				
MEDLINE	97044477						
COMMENT	Contact: Chin, H						
National Institute of Mental Health							
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD							
20892-9643, USA							
Tel: 301 443 1706							
Fax: 301 443 9890							
Email: mEST@mail.nih.gov							
Oligo-dt track not found, Not 1 site shown in beginning of sequence							

is likely internal to the message. cDNA library preparation: M.B. Soares Lab clone distribution: Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It should be noted that Bento Soares is generating a small number of additional specialized non-redundant arrays of BMAP cDNAs whose availability will be considered under appropriate and limited collaborative arrangements Seq primer: M13 Forward POLYA=No.

Location/Qualifiers

1. :450
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UI-M-BH3-awu-b-08-0-01"
 /clone_lib="NIH_BMAP_M_S4"
 /dev_stage="27-32 days"
 /lab_host="DH10B (Life Technologies)"

/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not 1; Site_2: Eco RI; The NIH_BMAP_M_S4 library is a subtracted library of a series, ultimately derived from a mixture of individually tagged normalized libraries from ten regions of the mouse brain (cerebellum, brain stems, olfactory bulbs, hypothalamus, cortex, amygdala, basal ganglia, pineal gland, striatum, hippocampus) after a series of subtractions to reduce the representation of cDNAs from which ESTs had already been generated. The following serially subtracted libraries were generated in this process: NIH_BMAP_M_S4, NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S3.1, NIH_BMAP_M_S2, NIH_BMAP_M_S1. The subtracted library (NIH_BMAP_M_S4) was constructed as follows: PCRamplified cDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived

NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S3.1, NIH_BMAP_M_S2, NIH_BMAP_M_S1. The subtracted library (NIH_BMAP_M_S4) was constructed as follows: PCRamplified cDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with a pool of the NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1 libraries in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Lifetechnologies) to generate the NIH_BMAP_M_S4 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)
 TAG_SEQ=None found"

6:791-806, 1996)

RESULTS 13

Query Match 30.8%; DB 12; Length 450;

Best Local Similarity 99.7%; Pred. No. 8.2e-59;

Matches 393; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

BASE COUNT

62 a 168 c 156 g 63 t 1 others

ORIGIN

Db	357	CGGGGCCGCACGCTGCAGGGTGCAGCCAGCCCCGACGCCATCGGGTC	416
OY	363	CTGTCAAACCCCGTGGCCACGCCGGCAGGGAGA	396
Db	417	CTGTCACCCCGTGGNCACGCCGGCGCCGGAGA	450
RESULT	14		
BB704449	BB704449	419 bp mRNA linear EST 11-OCT-2001	MUS
LOCUS		BB704449 RIKEN full-length enriched, in vitro fertilized eggs MUS	
DEFINITION		musculus cDNA clone 7420464A04 3', mRNA sequence.	
ACCESSION		BB704449	GI:16053284
VERSION		BB704449.1	
KEYWORDS		EST.	
SOURCE		house mouse.	
ORGANISM		Mus musculus	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE	1	(bases 1 to 419)	
AUTHORS	Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T., Ishii Hayatsu,N., Hiramoto,K., Hiraoaka,T., Hirozane,T., Imotani,K., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Watanuki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.		
TITLE	RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al. 2001)		
JOURNAL	Unpublished (2001)		
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.go.jp' URL:http://genome.gsc.riken.go.jp/' Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. <i>Genome Res.</i> . 10 (10), 1617-1630 (2000) wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanuki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. <i>Genome Res.</i> . 10 (11), 1757-1771 (2000) Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara Y., and Hayashizaki,Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. <i>Genome Res.</i> . 11 (2), 281-289 (2001) Please visit our web site (http://genome.gsc.riken.go.jp) for further details.		
FEATURES	e mouse tissues.		
Source	Location/Qualifiers		
1. .419	/organism="Mus musculus"		
/strain="C57BL/6J"	/clone="7420464A04"		
/db_xref="taxon:10090"	fertilized eggs"		
/sex="female"	/tissue_type="in vitro fertilized eggs"		
/dev_stage="egg"	/lab_host="DHL0B"		
/note="Site_1: Sali; Site_2: BamHI; cDNA library was			
RESULT	15		
BB703869	BB703869	521 bp mRNA linear EST 11-OCT-2001	MUS
LOCUS	BB703869	BB703869 RIKEN full-length enriched, in vitro fertilized eggs MUS	
DEFINITION		musculus cDNA clone 7420457C21 3', mRNA sequence.	
ACCESSION	BB703869		
VERSION	BB703869.1	GI:16052704	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE	1	(bases 1 to 521)	
AUTHORS	Hayatsu,N., Hiramoto,K., Hiraoaka,T., Hirozane,T., Imotani,K., Ishii Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T., Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,		

prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [^{5'} GAGAGAGAGAGATCCAAGAGCTCTTTPTTTTPTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [^{5'} GAGAGAGATCTCGAGTTAAATTAAATCCCCCCCC 3']. cDNA was cloned into the XbaI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC. 1. Cloning sites, 5' end: SalI; 3' end: BamHI"

BASE COUNT
ORIGIN
Query Match
Best Local Similarity
Matches 416; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 834 CAAGGACTGCAAATCCGGGGAGAGGCCCTATGTGTGGTGTGCAGGGCACCAGTA 893
Db 1 CAAGGACTGCAAATCCGGGGAGAGGCCCTATGTGTGGTGTGCAGGGCACCAGTA 60

QY 894 GGTGTTACTTCAAACAGTCTGGCAGTGTGTGAGAAATCCTACAAACCTTACAGAGTG 953
Db 61 GGT-TTACTTCAAACAGTCTGGCAGTTGTGAGAAATCCTACAAACCTTACAGAGTG 119

QY 954 AGGACATCACCTGTCAAAGTGTAAAGAACTAGATGTGCTGCCAGTCAGATTGCC 1013
Db 120 AGGACATCACCTGTCAAAGTGTAAAGAACTAGATGTGCTGCCAGTCAGATTGCC 179

QY 1014 ACGTGGACCTAAACGCCCATCGGCAGACTGTGTGGAGATGCAAGGACAACGCC 1073
Db 180 ACGTGGACCTAAACCCCCCATCGGCAGACTGTGTGGAGATGCAAGGACAACGCC 239

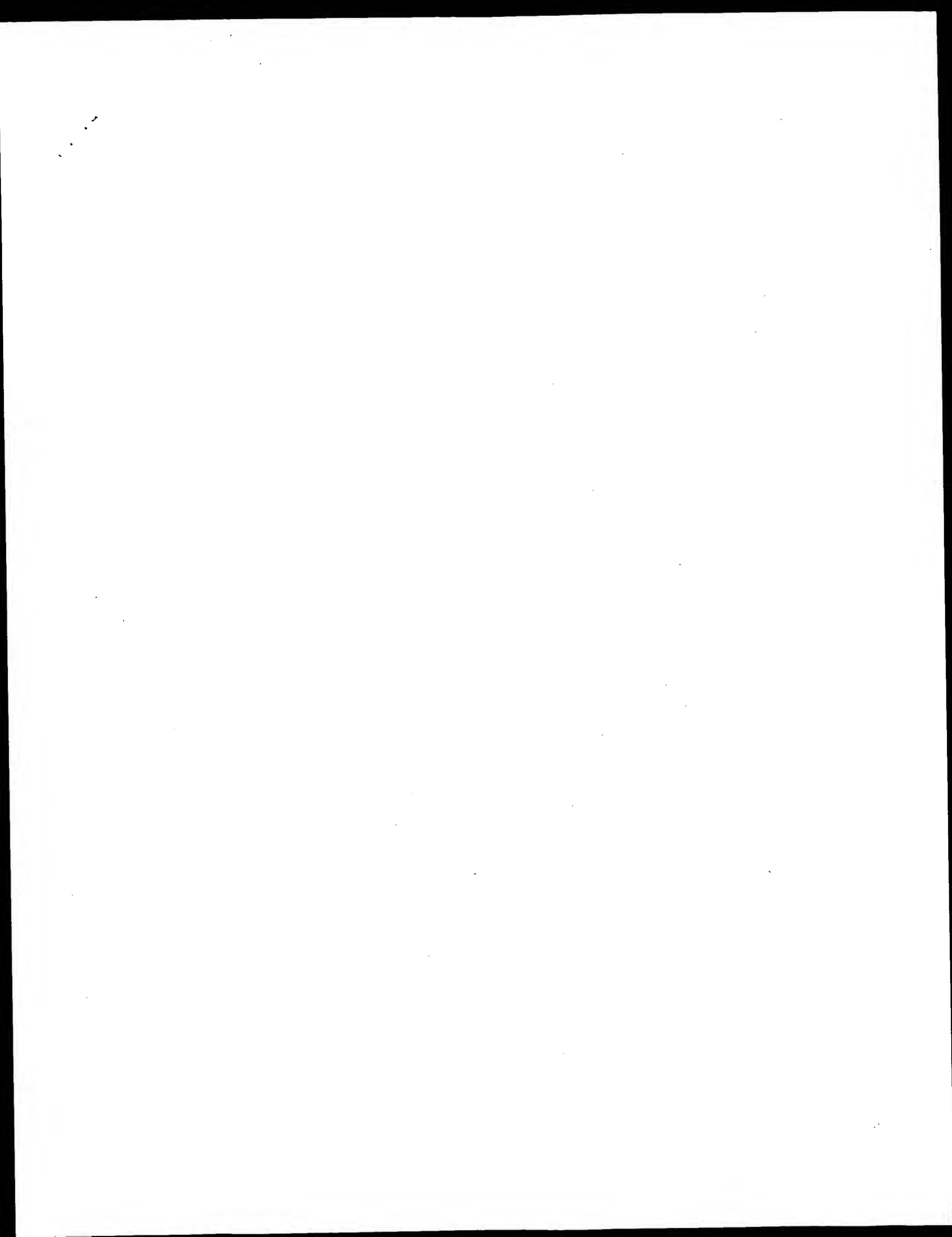
QY 1074 TGTCTCTGGCAGACCCCTCAGCTCAATACATCATTTAGTGAGTCAGACTTC 1133
Db 240 TGTCTCTGGCAGACCCCTCAGCTCAATACATCATTTAGTGAGTCAGACTTC 299

QY 1134 TGCTAGATGGGCTAATGGAATGGACAAGTGAGCTTCTCCCTCTCACCTCTCCCT 1193
Db 300 TGCTAGATGGGCTAATGGAATGGACAAGTGAGCTTCTCCCTCTCACCTCTCCCT 359

QY 1194 TCCAATTCATGACAGACAGTGTACTTGGATAAACGGTATT 1253
Db 360 TCCAATTCATGACAGACAGTGTACTTGGATAAACGGTATT 418

JOURNAL	Unpublished (2001)
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.go.jp URL: http://genome.gsc.riken.go.jp/
FEATURES	<p>Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.</p> <p>Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. <i>Genome Res.</i> 10 (10), 1617-1630 (2000)</p> <p>wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.</p> <p>RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. <i>Genome Res.</i> 10 (11), 1757-1771 (2000)</p> <p>Konno,H., Fukunishi,Y., shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.</p> <p>Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. <i>Genome Res.</i> 11 (2), 281-289 (2001)</p> <p>Please visit our web site (http://genome.gsc.riken.go.jp) for further details. mouse tissues.</p>
SOURCE	<p>Location/Qualifiers</p> <p>1. .521</p> <p>/organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="7420457C21" /clone_lib="RIKEN full-length enriched, in vitro fertilized eggs" /sex="female" /tissue_type="in vitro fertilized eggs" /dev_stage="egg" /lab_host="DH10B" /note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAGATCCAAAGAGCTCTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATCCCCCCCCCC 3']. cDNA was cloned into the XbaI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI"</p> <p>BASE COUNT</p> <p>153 a 138 c 119 g 111 t</p> <p>ORIGIN</p>
Query Match	28.8%; Score 367.4; DB 10; Length 521;
Best Local Similarity	87.3%;
Matches	448; Conservative 0; Mismatches 61; Indels 4; Gaps 4;
BASE COUNT	739
ORIGIN	12

Search completed: January 19, 2003, 20:41:16
Job time : 2230 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

OM nucleic - nucleic search, using sw model

Run on: January 19, 2003, 20:01:55 ; Search time 3366 Seconds
(without alignments)
11041.082 Million cell updates/sec

Title: US-09-844-864-1
Perfect score: 1277
Sequence: 1 aaggccggcgaggcgccgga.....acaaaaaaaaaaaaaaa 1277

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
 1: gb_ba:*
 2: gb_htg:*
 3: gb_in:*
 4: gb_om:*
 5: gb_ov:*
 6: gb_pat:*
 7: gb_ph:*
 8: gb_pl:*
 9: gb_pr:*
 10: gb_ro:*
 11: gb_sts:*
 12: gb_sy:*
 13: gb_uni:*
 14: gb_v1:*
 15: em_ba:*
 16: em_fun:*
 17: em_hum:*
 18: em_in:*
 19: em_mu:*
 20: em_om:*
 21: em_or:*
 22: em_ov:*
 23: em_pat:*
 24: em_ph:*
 25: em_pl:*
 26: em_ro:*
 27: em_sts:*
 28: em_un:*
 29: em_v1:*
 30: em_htg_hum:*
 31: em_htg_inv:*
 32: em_htg_other:*
 33: em_htg_mus:*
 34: em_htg_p1n:*
 35: em_htg_r1od:*
 36: em_htg_mam:*
 37: em_htg_vrt:*
 38: em_sy:*
 39: em_htgo_hum:*
 40: em_htgo_mus:*
 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

Result NO.	Score	Query	Match Length	DB	ID	Description
1	799.4	62.6	200535	10	AL671880	AL671880 Mouse DNA
2	689.8	54.0	144963	2	AC107686	AC107686 Mus muscu
3	657.4	51.5	67535	2	AC122733	AC122733 Mus muscu
4	615.4	48.2	173007	2	AC125993	AC125993 Rattus no
5	585.2	45.8	169551	2	AC127083	AC127083 Rattus no
6	452.4	35.4	169908	2	AC126519	AC126519 Rattus no
7	360	28.2	212848	2	AC108848	AC108848 Mus muscu
8	299.8	23.5	212848	2	AC108848	AC108848 Mus muscu
9	194.8	15.3	89551	2	AC127083	AC127083 Rattus no
10	147.6	11.6	138890	9	AC007970	AC007970 Homo sapi
11	147.6	11.6	151961	2	AC037443	AC037443 Homo sapi
12	147.6	11.6	204268	2	AC092163	AC092163 Homo sapi
13	134.6	10.5	187578	9	AC096952	AC096952 Homo sapi
14	73.2	5.7	694	9	HSA335296	AJ335296 Homo sapi
15	72.2	5.7	125020	9	AF429315	AF429315 Homo sapi
16	66.8	5.2	146431	2	AC109670	AC109670 Rattus no
17	64.4	5.0	172433	2	AC118973	AC118973 Rattus no
18	63.2	4.9	125020	9	AF429315	AF429315 Homo sapi
19	62.8	4.9	194873	2	AC125563	AC125563 Rattus no
20	62.6	4.9	1393	11	PM11H12G	AL684264 Penicill1
21	62	4.9	136551	2	AC048354	AC048354 Homo sapi
22	62	4.9	138467	2	AC111697	AC111697 Rattus no
23	59	4.6	185263	2	AC129765	AC129765 Rattus no
24	58.6	4.6	194936	2	AC115880	AC115880 Mus muscu
25	58.4	4.6	36394	1	SCD35	AL160312 Streptomy
26	58	4.5	92310	2	AC108380	AC108380 Pan trogl
27	58	4.5	97065	2	AC121735	AC121735 Rattus no
28	57.2	4.5	62847	2	AC125560	AC125560 Rattus no
29	57.2	4.5	83629	2	AC110485	AC110485 Rattus no
30	57.2	4.5	144053	2	AC092857	AC092857 Rattus no
31	57.2	4.5	150090	2	AC121750	AC121750 Rattus no
32	57.2	4.5	150238	2	CNS08CA9	AL831796 Oryza sat
33	57	4.5	92055	2	AC105692	AC105692 Rattus no
34	57	4.5	115553	2	AC126077	AC126077 Rattus no
35	56.8	4.4	73041	2	AC127648	AC127648 Rattus no
36	56.8	4.4	146431	2	AC109670	AC109670 Rattus no
37	56.6	4.4	157448	2	AC131139	AC131139 Rattus no
38	56.4	4.4	136788	2	AC117044	AC117044 Rattus no
39	56.2	4.4	159677	2	AC113802	AC113802 Rattus no
40	56.2	4.4	184402	2	AC127041	AC127041 Rattus no
41	55.8	4.4	179608	2	AC128497	AC128497 Rattus no
42	55.6	4.4	991	11	PM12H12B	AL684455 Penicill1
43	55.6	4.4	134366	2	AC119661	AC119661 Rattus no
44	55.6	4.4	136014	2	AC126531	AC126531 Rattus no
45	55.6	4.4	151349	10	AL355176	AL355176 Mouse DNA

ALIGNMENTS

RESULT 1
AL671880
LOCUS AL671880 200535 bp DNA linear R09-AUG-2002
DEFINITION Mouse DNA sequence from clone RP23-384C22 on chromosome X, complete
sequence.
ACCESSION AL671880
VERSION AL671880.14 GI:22204349
KEYWORDS HTG.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 200535)
AUTHORS Van Hellmond,Z.
TITLE Direct Submission

AUTHORS

Birren, B., Linton, L., Nusbaum, C., Landre, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boquslavkiy, L., Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyste, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lehoczky, J., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talama, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainon, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

JOURNAL

Submitted (21-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On Apr 21, 2002 this sequence version replaced gi:18308388. All repeats were identified using RepeatMasker:

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Project name: L17749

Center clone name: 228_B_12

Summary Statistics

sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 141008 bases at least 040

Consensus quality: 142434 bases at least 030

Consensus quality: 142970 bases at least 020

Insert size: 140000; agarose-fp

Insert size: 143263; sum-of-contigs

Quality coverage: 8.2 in Q20 bases; agarose-fp

Quality coverage: 8.0 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 18 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter.

This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.

- * 1 630: contig of 630 bp in length
- * 631 730: gap of 100 bp
- * 731 2440: contig of 1710 bp in length
- * 2441 2540: gap of 100 bp
- * 2541 4613: contig of 2073 bp in length
- * 4614 4713: gap of 100 bp
- * 4714 6263: contig of 1550 bp in length
- * 6264 6363: gap of 100 bp
- * 6364 9200: contig of 2837 bp in length
- * 9201 9300: gap of 100 bp
- * 9301 12704: contig of 3404 bp in length
- * 12705 12804: gap of 100 bp
- * 12805 15990: contig of 3186 bp in length
- * 15991 16090: gap of 100 bp
- * 16091 20096: contig of 4006 bp in length

FEATURES

source

1. 144963

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="RP24-228B12"

/clone_1lib="RPCI-24 Male Mouse BAC"

1. .630

/note="assembly_fragment"

731. .2440

/note="assembly_fragment"

2541. .4613

/note="assembly_fragment"

4714. .6263

/note="assembly_fragment"

6364. .9200

/note="assembly_fragment"

9301. .12704

/note="assembly_fragment"

12805. .15990

/note="assembly_fragment"

16091. .20096

/note="assembly_fragment"

20197. .23757

/note="assembly_fragment"

23858. .35345

/note="assembly_fragment"

35446. .43937

/note="assembly_fragment"

44038. .54685

/note="assembly_fragment"

54786. .64286

/note="assembly_fragment"

64387. .77535

/note="assembly_fragment"

77536. .77635

/note="assembly_fragment"

91483. .91582

/note="assembly_fragment"

103491. .124554

/note="assembly_fragment"

124555. .124654

/note="assembly_fragment"

124655. .144963

/note="assembly_fragment"

144963: contig of 20309 bp in length.

Location/Qualifiers

1. 144963

BASE COUNT

ORIGIN

Query	Match	Score	DB	Length
QY	3	ggggggaggccggccgaccatgttccggcgagcacgttccacccctccggca	62	144963;
Db	142184	ggggggaggccggcaacgcacccatgttccggcgagcacgttccacccctccggca	62	142243

QY	63	TCCTTATCCGCAGGCCACCAAGCGGGATGGCTGGAGCAGGGCTGCCG	122
Db	142244	TCTTATCCGCAGGCCACCAAGCGGGATGGCTGGAGCAGGGCTGCCG	142303
QY	123	ACCCGAGCCCTCCTCCCGGCTACAGACAGCTCATGGCCGGAGTACGTCA	182
Db	142304	ACCCGGCCCTCTCCCTCCCGGCTACAGACAGCTCATGGCCGGAGTACGTCA	142363
QY	183	CAGGCCACCAAGGGCACAGCTCATGGCCCTGCTCGGGATGGTCAG	242
Db	142364	CGGGCCACCAAGGGCACAGCTATGGCCCTGCTCGGGATGGTCAG	142423
QY	243	CAGCCGTGACGCTGCGGTGAGGTGAACCCGGCGACGCCCTCGGTGAGT	302
Db	142424	CGGGCCACCAAGGGCACAGCTATGGCCCTGCTCGGGATGGTCAG	142483
QY	303	CGGGCCACCAAGGGCACAGCTCATGGCCGGAGTACGTGTTCA	362
Db	142484	CGGGGCCACGCTGCAGCCTGCAGGGTGCCGAGGCCAGCCCCGACGCC	142543
QY	363	CTGTCACACCCCGTGGCACGCCACGCCGGGGAGATCCCCGCGATCCTGGCAGACCGTAGC	422
Db	142544	CTGTCACACCCCGTGGCACGCCACGCCGGGGAGATCCCCGCGATCCTGGCAGACCGTAGC	142603
QY	423	CCCGTCTCGTCCGTGACCTCTGTGGCCTCTCCTCACTGGAGGTGCGGGAG	482
Db	142604	CCCGTCTCGTCCGTGACCTCTGTGGCCTCTCCTCACTGGAGGTGCGGGAG	142658
QY	483	GCAGACACCCACGAAAGGGAGAGGGAGGCCATCTGGGACCCGGAACCGGAGC	542
Db	142659	GCAGACACCCACGAAAGGGAGAGGGAGGCCATCTGGGACCCGGAACCGGAGC	142718
QY	543	GAGAGAGTGGCCGAGGAAAGCGGTCCCCCAGCCCGAAGCGAGGGCGATGTTCA	602
Db	142719	GAGAGAGTGGCCGAGGAAAGCGGTCCCCCAGCCCGAAGCGAGGGCGACGTCCA	142778
QY	603	GGCTGCAGGCCAGGCCGGGGAGCAGCAGCCACCCGGAGGACCGGAACAGTGTGCC	662
Db	142779	GGCTGAAGGGCAGGCCGGGGAGCAGCCACCCGGAGGACCGGAACAGTGTGCC	142838
QY	663	GGCGATGCACTGTGAGCCTGGGAGGCCATGTCTGCCGCAGAGATGGCTCAGGA	722
Db	142839	GGCGATGCACTGTGAGCCTGGGAGGCCATGTCTGCCGCAGAGATGGCTCAGGA	142898
QY	723	CCCCGGTATTGGATGCCCTCGAGACCAGGCCAAAGCACGGAGCAGGACAA	782
Db	142899	CCCCGGTACTCGGCTGGAGTGTGAGATGGCTCAGGA	142945
QY	783	GGAGCGCTGGTTCCAGTT 803	
Db	142946	GGAGCTCTGGTTCCAGTT 142966	
RESULT	3		
AC122733			
LOCUS	AC122733	67535 bp	DNA linear HTG 25-MAY-2002
DEFINITION	Mus musculus clone RP24-506B15, LOW-PASS SEQUENCE SAMPLING.		
ACCESSION	AC122733		
VERSION	AC122733.1	GI:21206317	
KEYWORDS	HTG; HTGS_PHASE0.		
SOURCE	Mus musculus.		
ORGANISM	Mus musculus		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE	1 (bases 1 to 67535)		
AUTHORS	Birren, B., Linton, L., Nusbaum, C. and Lander, E.		
TITLE	Mus musculus, clone RP24-506B15		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 67535)		
AUTHORS	Birren, B., Linton, L., Nusbaum, C., Landier, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,		

Query	Match	Score	DB	Length	Best Local Similarity	Pred.	No.	Mismatches	Indels	Gaps
Qy 23985	24084: gap of unknown length	*	133313	138145: contig of 4833 bp in length	*	*	138245	gap of unknown length		
24085	25116: contig of 1032 bp in length	*	138246	138245: gap of unknown length	*	*	143564	contig of 5319 bp in length		
25117	25216: gap of unknown length	*	138246	143564: contig of 5319 bp in length	*	*	143565	gap of unknown length		
25217	27027: contig of 1811 bp in length	*								
27028	27127: gap of unknown length	*								
27128	29170: contig of 2043 bp in length	*								
29171	29270: gap of unknown length	*								
29271	31495: contig of 2225 bp in length	*								
31495	31595: gap of unknown length	*								
31596	34124: contig of 2529 bp in length	*								
34125	34224: gap of unknown length	*								
34225	37124: contig of 2900 bp in length	*								
37125	37224: gap of unknown length	*								
37225	39707: contig of 1684 bp in length	*								
39708	39807: gap of unknown length	*								
39808	41454: contig of 1281 bp in length	*								
41455	41554: gap of unknown length	*								
41555	43239: contig of 1684 bp in length	*								
43239	43338: gap of unknown length	*								
43339	44619: contig of 1647 bp in length	*								
44620	44719: gap of unknown length	*								
44720	46704: contig of 1985 bp in length	*								
46705	46804: gap of unknown length	*								
46805	49282: contig of 2478 bp in length	*								
49283	49383: gap of unknown length	*								
51920	52019: gap of unknown length	*								
52020	54858: contig of 2839 bp in length	*								
54859	54959: gap of unknown length	*								
57482	57581: gap of unknown length	*								
57582	59384: contig of 1803 bp in length	*								
59385	59484: gap of unknown length	*								
59485	60768: contig of 1284 bp in length	*								
60769	60868: gap of unknown length	*								
60869	62575: contig of 1707 bp in length	*								
62576	62675: gap of unknown length	*								
62676	64968: contig of 2293 bp in length	*								
64969	65068: gap of unknown length	*								
65069	68642: contig of 3574 bp in length	*								
68643	68742: gap of unknown length	*								
68743	71330: contig of 2588 bp in length	*								
71331	71431: gap of unknown length	*								
71431	75416: contig of 3985 bp in length	*								
75516	75515: gap of unknown length	*								
78192	78191: contig of 2676 bp in length	*								
78192	78291: gap of unknown length	*								
78292	82095: contig of 3804 bp in length	*								
82096	82196: gap of unknown length	*								
82196	86303: contig of 4107 bp in length	*								
86303	86402: gap of unknown length	*								
86403	89944: contig of 3542 bp in length	*								
89945	90044: gap of unknown length	*								
90045	94438: contig of 4393 bp in length	*								
94438	94538: gap of unknown length	*								
94538	97928: contig of 3391 bp in length	*								
97929	98028: gap of unknown length	*								
98029	101684: contig of 3656 bp in length	*								
101685	101785: gap of unknown length	*								
101785	105383: contig of 3599 bp in length	*								
105384	105483: gap of unknown length	*								
105484	108194: contig of 2711 bp in length	*								
108195	108294: gap of unknown length	*								
108295	111907: contig of 3613 bp in length	*								
111908	112008: gap of unknown length	*								
112008	117332: contig of 5225 bp in length	*								
117333	121679: gap of unknown length	*								
121680	121780: contig of 5836 bp in length	*								
121780	127616: gap of unknown length	*								
127616	133212: contig of 5497 bp in length	*								
133213	133312: gap of unknown length	*								

RESULT 5
AC127083/c
LOCUS AC127083 Rattus norvegicus clone CH230-69F8, *** SEQUENCING IN PROGRESS ***,
DEFINITION 48 unordered pieces.
ACCESSION AC127083
VERSION AC127083.1 GI:21743720

KEYWORDS	HTG; HTGS_PHASE1.
SOURCE	Norway rat.
ORGANISM	Rattus norvegicus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
AUTHORS	1 (bases 1 to 89551)
MUZNY, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Alsbrooks, S.L., Amaralunge, H.C., Are, J.R., Ayele, M., Banks, T., Barbaria, J., Benton, J., Brimage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhinney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newtson, J., Newtson, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoshtari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telfrod, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstock, G., and Gibbs, R.	
TITLE	Direct Submission
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 89551)
AUTHORS	Worley, K.C.
TITLE	Direct Submission
JOURNAL	Submitted (13-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
REFERENCE	3 (bases 1 to 89551)
AUTHORS	Worley, K.C.
TITLE	Direct Submission
JOURNAL	Submitted (31-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
COMMENT	Center: Baylor College of Medicine Center code: BCM Center: Baylor College of Medicine Center: Baylor College of Medicine Center: Baylor College of Medicine Web site: http://www.hgsc.bcm.tmc.edu/ Contact: hgsc-help@bcm.tmc.edu ----- Project Information Center project name: KAMV Center clone name: CH230-69F8 ----- Summary Statistics Sequencing vector: Plasmid: Chemistry: Dye-terminator Big Dye: 100% of reads Assembly program: Phrap; version 0.990329

KEYWORDS: HTG; HTGS_PHASE1.
 SOURCE: Norway rat.
 ORGANISM: Rattus norvegicus
 REFERENCE: Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

AUTHORS: 1 (bases 1 to 89551)
 MUZNY, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Alsbrooks, S.L., Amaralunge, H.C., Are, J.R., Ayele, M., Banks, T., Barbaria, J., Benton, J., Brimage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhinney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newtson, J., Newtson, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoshtari, N., Sisson, I., Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telfrod, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstock, G., and Gibbs, R.

KEYWORDS: Consensus quality: 42669 bases at least 040
 SOURCE: Consensus quality: 46367 bases at least 030
 ORGANISM: Consensus quality: 48722 bases at least 020

REFERENCE: (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 AUTHORS: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 AUTHORS: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

REFERENCE: * NOTE: This is a 'working draft' sequence. It currently consists of 48 contigs. The true order of the pieces is not

QY	466	GAGGTTGGGGAGGCAGGCCAGACCCCACGAAGGGAGAGGGAGAGGGAGACCGGGCATCTCGGG	46667	gap of unknown length	Db	39327	TCCTGGGCACCGTGCCTGTACTCGCCCGTGCACCTTCGGTGGCCTCTCCTCGCTG	46668
QY	49450	49450: gap of unknown length	49449	contig of 2782 bp in length	Db	39267	GAGGTGGGGAGACAGGCCAACGAAAGGGAGAGGGAGACCGGGCACCGGGCAAGC	49549
QY	51783	51783: gap of unknown length	51782	contig of 2036 bp in length	Db	54046	54046: contig of 2163 bp in length	51784
QY	56183	56183: gap of unknown length	56282	contig of 1652 bp in length	Db	54147	54147: gap of unknown length	56184
QY	56283	56283: gap of unknown length	57934	contig of 1610 bp in length	Db	54147	54147: contig of 2036 bp in length	56183
QY	57935	57935: gap of unknown length	58034	contig of 1791 bp in length	Db	57935	57935: gap of unknown length	57935
QY	58035	58035: gap of unknown length	59826	contig of 1838 bp in length	Db	58035	58035: gap of unknown length	58035
QY	59826	59826: gap of unknown length	59925	contig of 2165 bp in length	Db	59826	59826: gap of unknown length	59826
QY	59925	59925: gap of unknown length	61535	contig of 2803 bp in length	Db	59925	59925: gap of unknown length	59925
QY	61535	61535: gap of unknown length	61635	contig of 2152 bp in length	Db	61535	61535: gap of unknown length	61535
QY	61635	61635: gap of unknown length	63474	contig of 1439 bp in length	Db	61635	61635: gap of unknown length	61635
QY	63474	63474: gap of unknown length	63574	contig of 1385 bp in length	Db	63474	63474: gap of unknown length	63574
QY	63574	63574: gap of unknown length	66376	contig of 3120 bp in length	Db	63574	63574: gap of unknown length	66376
QY	66376	66376: gap of unknown length	70180	contig of 1385 bp in length	Db	66376	66376: gap of unknown length	70180
QY	70181	70181: gap of unknown length	70280	contig of 2152 bp in length	Db	70181	70181: gap of unknown length	70280
QY	70281	70281: gap of unknown length	72432	contig of 3120 bp in length	Db	70281	70281: gap of unknown length	72432
QY	72433	72433: gap of unknown length	72532	contig of 1385 bp in length	Db	72433	72433: gap of unknown length	72532
QY	72533	72533: gap of unknown length	73918	contig of 1385 bp in length	Db	72533	72533: gap of unknown length	73918
QY	73918	73918: gap of unknown length	74018	contig of 3120 bp in length	Db	73918	73918: gap of unknown length	74018
QY	74018	74018: gap of unknown length	77138	contig of 1385 bp in length	Db	74018	74018: gap of unknown length	77138
QY	77138	77138: gap of unknown length	77238	contig of 2760 bp in length	Db	77138	77138: gap of unknown length	77238
QY	77238	77238: gap of unknown length	79998	contig of 2762 bp in length	Db	79998	79998: gap of unknown length	79998
QY	80098	80098: gap of unknown length	82859	contig of 2721 bp in length	Db	80098	80098: gap of unknown length	82859
QY	82860	82860: gap of unknown length	82960	contig of 2721 bp in length	Db	82860	82860: gap of unknown length	82960
QY	82960	82960: gap of unknown length	85680	contig of 3120 bp in length	Db	82960	82960: gap of unknown length	85680
QY	85680	85680: gap of unknown length	85780	contig of 3771 bp in length.	Db	85680	85680: gap of unknown length	85780
FEATURES		Location/Qualifiers						
SOURCE		1. .89551 /organism="Rattus norvegicus" /db_xref="taxon:10116" /clone="CH230-69F8"						
BASE COUNT	23140	a	19260	c	19626	g	22768	t
ORIGIN								
REFERENCE								
AUTHORS								
Query Match	45..8%	Score 585.2; DB 2; Length 89551;						
Best: Local Similarity 87.6%; Pred: No. 2.1e-110; Matches 664; Conservative 0; Mismatches 88; Indels 6; Gaps 2;								
QY	49	49 CACCCCTGCCGCATCCTTATCCGC--AGGCCACCAAAGCCGGGATGGCTGGAGGTC	105					
Db	39687	CACCCATGCCGCATCCTTACCCGCCCTCTGCAGCCAAGCCGGGATGGCTGGAGGTT	39628					
QY	106	106 GGAGCCAGGGCTGCCGACCCCCGGCCCTCTCTCCCCGCTACAGACAGCTCATG	165					
Db	39627	GGAGCCAGGGCTGCAGGCCCTCTCCCCGCTACAGACAGCTCATGGCCTGTCGCGGATG	39568					
QY	166	166 GCCGGGGAGTAGTCGACAGCCACCAGGGCACAGCTCATGGCCTGTCGCGGATG	225					
Db	39567	GCCGGGGAGTAGTCGACAGCTATCAGGGAGCCCTCATGGCCTGTCGCGAATG	39508					
QY	226	226 GGTCCCCGGTCGGTCAGCAGCCGCTGGGGTGCAGGTGAAACCCGGCCGAGGCC	285					
Db	39507	GGTCCCCGGTCGGTCAGCAGCCGACGCTGGGGTGCAGGTGAAACCCGGCCGAGGCC	39448					
QY	286	286 TCGGTGCAGGTGTCAGTCGGGCCACGGCTGCAGGGTGCAGGCCAGGCC	345					
Db	39447	TCGGTGCAGTGTCGGCTCGGCCACACTGCAGGCCCTGGACGGCCGAGGCC	39388					
QY	346	346 GACGGCCGATCGGTCTCGTCACCGCCGACGGCTGCAGGGTGCAGGCC	405					
Db	39387	39387 GACGGCCGCTGGTRCTCGCCAAACCCCCGAGCCCCGGAGACCCCCGGCA	39328					
QY	406	406 TCCTGGCAGACCGTAGCCCCGTCGTCGCGACTCTCGTGGCCTCTCCTCACTG	465					

Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I., *
 Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H., *
 Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., *
 Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., *
 Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., *
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., *
 Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K., *
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., *
 Weinstock,G. and Gibbs,R.

TITLE Direct Submission
 JOURNAL unpublished
 REFERENCE 2 (bases 1 to 169908)
 AUTHORS Worley,K.C.
 TITLE Direct Submission
 JOURNAL Submitted (06-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 REFERENCE 3 (bases 1 to 169908)
 AUTHORS Worley,K.C.
 TITLE Direct Submission
 JOURNAL Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 COMMENT Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu/
 ----- Project Information
 Center clone name: CH230-159N5
 ----- Summary Statistics
 sequencing vector: Plasmid;
 chemistry: dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap; version 0.990329
 consensus quality: 125808 bases at least Q40
 consensus quality: 132708 bases at least Q30
 consensus quality: 136844 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 60 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 1574: contig of 1574 bp in length
 * 1575 1674: gap of unknown length
 * 1675 2679: contig of 1005 bp in length
 * 2680 2779: gap of unknown length
 * 2780 3878: contig of 1099 bp in length
 * 3879 3978: gap of unknown length
 * 3979 5423: contig of 1445 bp in length
 * 5424 5523: gap of unknown length
 * 5524 6788: contig of 1265 bp in length
 * 6789 6888: gap of unknown length
 * 6889 8285: contig of 1397 bp in length
 * 8286 8385: gap of unknown length
 * 8386 9590: contig of 1205 bp in length
 * 9591 9690: gap of unknown length
 * 9691 10999: contig of 1309 bp in length
 * 11000 11099: gap of unknown length
 * 11100 12498: contig of 1399 bp in length
 * 12499 12598: gap of unknown length
 * 12599 14147: contig of 1549 bp in length
 * 14148 14247: gap of unknown length
 * 14248 16266: contig of 2019 bp in length
 * 16267 16366: gap of unknown length
 * 16367 18233: contig of 1867 bp in length
 * 18234 18333: gap of unknown length
 18334 19762 19761: contig of 1428 bp in length
 19862 19861: gap of unknown length
 21637 21638: contig of 1776 bp in length
 21738 21737: gap of unknown length
 23143 23144: contig of 1406 bp in length
 23243 23244: gap of unknown length
 25296 25297: contig of 2053 bp in length
 25396 25397: gap of unknown length
 26533 26534: contig of 1137 bp in length
 26633 26634: gap of unknown length
 28678 28278: contig of 1645 bp in length
 28279 28379: gap of unknown length
 30343 30344: contig of 1965 bp in length
 32122 32121: gap of unknown length
 32221 32222: contig of 1678 bp in length
 34582 34583: gap of unknown length
 34682 34683: contig of 1298 bp in length
 35980 35981: gap of unknown length
 36080 36081: contig of 1314 bp in length
 37394 37395: gap of unknown length
 37494 37495: contig of 1474 bp in length
 38968 38969: gap of unknown length
 39068 39069: contig of 1131 bp in length
 40200 40299: gap of unknown length
 40300 41804: contig of 1505 bp in length
 41805 41905: gap of unknown length
 43569 43570: contig of 1665 bp in length
 43669 43670: gap of unknown length
 45680 45681: contig of 2011 bp in length
 45780 45781: gap of unknown length
 48215 48216: contig of 2435 bp in length
 51068 51069: gap of unknown length
 51168 51169: contig of 2753 bp in length
 53666 53667: gap of unknown length
 53766 53767: contig of 2498 bp in length
 55832 55833: gap of unknown length
 55933 55934: contig of 2066 bp in length
 57560 57561: gap of unknown length
 57660 57661: contig of 1628 bp in length
 59904 59905: gap of unknown length
 60004 60005: contig of 2244 bp in length
 62974 62975: gap of unknown length
 63074 63075: contig of 2970 bp in length
 66491 66492: contig of 3417 bp in length
 66591 66592: gap of unknown length
 68271 68272: contig of 1680 bp in length
 68371 68372: gap of unknown length
 70762 70763: contig of 2391 bp in length
 70862 70863: gap of unknown length
 73263 73363: contig of 2401 bp in length
 73364 73364: gap of unknown length
 76903 76904: contig of 3540 bp in length
 77003 77004: gap of unknown length
 79444 79544: contig of 2441 bp in length
 81957 81958: contig of 2413 bp in length
 82057 82058: gap of unknown length
 85214 85215: contig of 3157 bp in length
 85314 85315: gap of unknown length
 87778 87779: contig of 2464 bp in length
 87878 87879: gap of unknown length
 90897 90898: contig of 3019 bp in length
 90997 90998: gap of unknown length
 96387 96388: contig of 5390 bp in length
 96487 96488: gap of unknown length
 100595 100596: contig of 4108 bp in length
 104991 104992: gap of unknown length
 105091 105092: contig of 4295 bp in length

	REFERENCE
AC108848	1 (bases 1 to 212848)
LOCUS	109487 113734: contig of 4248 bp in length
DEFINITION	Mus musculus clone RP23-300K5, WORKING DRAFT SEQUENCE, 36 ordered pieces.
ACCESSION	113835 116884: contig of 3050 bp in length
VERSION	116885 116984: gap of unknown length
KEYWORDS	116985 121008: contig of 4024 bp in length
SOURCE	121009 121108: gap of unknown length
ORGANISM	121109 124634: contig of 3526 bp in length
Mus musculus	124635 124734: gap of unknown length
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	124735 130307: contig of 5573 bp in length
RESULT	130308 130407: gap of unknown length
7	
AC108848	Query Match 35.4%; Score 452.4; DB 2; Length 169908;
Best Local Similarity 87.5%; Pred. No. 4.9e-83; Matches 530; Conservative 0; Mismatches 71; Indels 5; Gaps 3;	
OY 199 CAGCTCATGGCCCTGCTGCGGGATGGTCCCCGTC-GTCAGCAGCGTGACCGCTGC 257	
Db 10971 CAGCTCATGGCCTGCTGCGGAATGGTCCCCGTGCAGCAGCGCGACGCTGC 10912	
OY 258 GGTGAGGTGAAACCCGCCGACGCCCTCGGTGCACTCGGCGCCGACCGCT 317	
Db 10911 GGTGAGGTGAAACCCGCCGCGATGCTCGCTCGCTCGCTCGGCGCACACT 10852	
OY 318 GCAGCCTGAGGTGCCGAGCCAGCCCCGACGCCGATCGGGTCTGTCAACCCCGTGG 377	
Db 10851 GCAGCCTGGACGCCGAGCCAGCCCCGACGCCGAGCCCCCTGGTCTGCCAACCCCGAG 10792	
OY 378 CCAGCGCCGGGGAGATCCCCGGATCTGGAGACCGTAGGCCCGTCTCGTCCGT 437	
Db 10791 CCCCGCCAGGGGGAGACCCCGCATCTGGGCACCGCTGCCCTGACTCGCCCGT 10732	
OY 438 GACCTTCTGTGCCCTCTCTCTACTGGAGTTGGGGAGGCAGACACCCACGAA 497	
Db 10731 GACCTTGGTGGCCTCTCTCGCTGGAGGTGGGGAGACAGGCCACGAA 10672	
OY 498 GGGAGAGGGAGCCCGGCATCCCTCGGGACCCGGAGGGAGAGAGGGTGGCCG 557	
Db 10671 -GGGAGGGAGACCCGGCACCCACGGGGACCCCCGGAGCCGGAGCTGCAGT 10613	
OY 558 GAGGAAGGGTCCCCCAGGGGAAGCGGAGGGGAGATGTTCAGGCTGCAGGGCAGG 617	
Db 10612 GATGAAAGCAGTCCCCAGGGAGAGCCGGAGCTGCAGGCTGAAGGGCAGGA 10553	
OY 618 CGGGTGGAGCAGGCCACCACGGAGGACCGGGAAACAGTGTGGGGGATGCTGA 677	
Db 10552 TG--GGCAGGAGCAGCCACGGGGAGGACAGTGTGGGGCATGCAGTCTGA 10496	
OY 678 GCCTGGAGGCCATGTGTGGGGAGAGATGGCTCAGGAGCCCCGGTGTGCA 737	
Db 10495 GCCCCGGAGTGGAGGCCACCTCTGTGTGAGATGGCTCAGGACCCCAGTGCG 10436	
OY 738 TGCCCCCTCGAGACCAAGGGCTCCCGCAAAGCACGGAGGACAAGGGAGGCTGCGTT 797	
Db 10435 TGCCTCTAGAGACCGGGCTCCCCACAGAGCACTGAGCAGCAAGGGAGGCCCTGCGTT 10376	
OY 798 CCAGTT 803	
Db 10375 CCAGGT 10370	
COMMENT	
TITLE	
JOURNAL	
REDFERENCES	
AUTHORS	
Birren,B., Linton,L., Nusbaum,C., and Lander,E.	
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Macdonald,P., Major,J., Marquis,N., Matthews,C., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Miheva,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Rettar,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Watson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.	
TITLE	
JOURNAL	
REDFERENCES	
AUTHORS	
Birren,B., Linton,L., Nusbaum,C., and Lander,E.	
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Macdonald,P., Major,J., Marquis,N., Matthews,C., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Miheva,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Rettar,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.	
TITLE	
JOURNAL	
REDFERENCES	
AUTHORS	
Birren,B., Linton,L., Nusbaum,C., and Lander,E.	
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Macdonald,P., Major,J., Marquis,N., Matthews,C., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Miheva,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Rettar,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.	
TITLE	
JOURNAL	
REDFERENCES	
AUTHORS	
Birren,B., Linton,L., Nusbaum,C., and Lander,E.	
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Macdonald,P., Major,J., Marquis,N., Matthews,C., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Miheva,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Rettar,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.	
TITLE	
JOURNAL	
REDFERENCES	
AUTHORS	
Birren,B., Linton,L., Nusbaum,C., and Lander,E.	
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Macdonald,P., Major,J., Marquis,N., Matthews,C., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Miheva,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Rettar,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.	
TITLE	
JOURNAL	
REDFERENCES	
AUTHORS	
Birren,B., Linton,L., Nusbaum,C., and Lander,E.	
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Macdonald,P., Major,J., Marquis,N., Matthews,C., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Miheva,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Rettar,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.	
TITLE	
JOURNAL	
REDFERENCES	
AUTHORS	
Birren,B., Linton,L., Nusbaum,C., and Lander,E.	
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Macdonald,P., Major,J., Marquis,N., Matthews,C., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Miheva,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Rettar,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.	
TITLE	
JOURNAL	
REDFERENCES	
AUTHORS	
Birren,B., Linton,L., Nusbaum,C., and Lander,E.	
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Macdonald,P., Major,J., Marquis,N., Matthews,C., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Miheva,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Rettar,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.	
TITLE	
JOURNAL	
REDFERENCES	
AUTHORS	
Birren,B., Linton,L., Nusbaum,C., and Lander,E.	
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Macdonald,P., Major,J., Marquis,N., Matthews,C., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Miheva,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Rettar,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.	
TITLE	
JOURNAL	
REDFERENCES	
AUTHORS	
Birren,B., Linton,L., Nusbaum,C., and Lander,E.	
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Macdonald,P., Major,J., Marquis,N., Matthews,C., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Miheva,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Rettar,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.	
TITLE	
JOURNAL	
REDFERENCES	
AUTHORS	
Birren,B., Linton,L., Nusbaum,C., and Lander,E.	
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Macdonald,P., Major,J., Marquis,N., Matthews,C., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Miheva,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Rettar,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.	
TITLE	
JOURNAL	
REDFERENCES	
AUTHORS	
Birren,B., Linton,L., Nusbaum,C., and Lander,E.	
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Macdonald,P., Major,J., Marquis,N., Matthews,C., McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Miheva,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Rettar,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J	

Consensus quality: 206302 bases at least 030
 Consensus quality: 208218 bases at least Q20
 Insert size: 209348; sum-of-contigs
 Quality coverage: 5.7 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 36 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submittor.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

FEATURES	source
1. .212848	Location/Qualifiers
/organism="Mus musculus"	
/db_xref="taxon:10090"	
/clone="RP23-300K5"	
/clone_lib="RPCI-23 Female Mouse BAC"	
1. .641	
/note="assembly_fragment"	
742. .1427	
/note="assembly_fragment"	
1428 1527: gap of 100 bp	
1528 3170: contig of 1643 bp in length	
3171 3270: gap of 686 bp in length	
3271 4396: contig of 1126 bp in length	
4397 4496: gap of 100 bp	
4497 5909: contig of 1413 bp in length	
5910 6009: gap of 100 bp	
6010 7583: contig of 1574 bp in length	
7584 7683: gap of 100 bp	
7684 8975: contig of 1292 bp in length	
8976 9075: gap of 100 bp	
9076 10257: contig of 1182 bp in length	
10258 10357: gap of 100 bp	
10358 11608: contig of 1251 bp in length	
11609 11708: gap of 100 bp	
11709 13289: contig of 1581 bp in length	
13290 13389: gap of 100 bp	
13390 15335: contig of 1946 bp in length	
15336 15435: gap of 100 bp	
15436 16250: contig of 815 bp in length	
16251 16350: gap of 100 bp	
16351 17669: contig of 1319 bp in length	
17670 1769: gap of 100 bp	
17770 19907: contig of 2138 bp in length	
19908 20007: gap of 100 bp	
20008 21470: contig of 1463 bp in length	
21471 21570: gap of 100 bp	
21571 23086: contig of 1516 bp in length	
23087 23186: gap of 100 bp	
23187 24723: contig of 1537 bp in length	
24724 24823: gap of 100 bp	
24824 24823: gap of 100 bp	
27747 27846: contig of 2923 bp in length	
27847 29873: contig of 2027 bp in length	
29874 29973: gap of 100 bp	
29974 33212: contig of 3239 bp in length	
33213 33312: gap of 100 bp	
33313 61732: contig of 28420 bp in length	
61733 61832: gap of 100 bp	
61833 63792: contig of 1960 bp in length	
63793 63892: gap of 100 bp	
63893 67231: contig of 3339 bp in length	
67232 67331: gap of 100 bp	
67332 70922: contig of 3591 bp in length	
70923 71022: gap of 100 bp	
71023 76930: contig of 5908 bp in length	
76931 77030: gap of 100 bp	
77031 82623: contig of 5593 bp in length	
82624 82723: gap of 100 bp	
82724 87875: contig of 5152 bp in length	
87876 87975: gap of 100 bp	
87976 93774: contig of 5799 bp in length	
93775 93874: gap of 100 bp	
93875 104205: contig of 10331 bp in length	
104206 104305: gap of 100 bp	
104306 119104: contig of 14799 bp in length	

Query Match 28.2%; Score 360; DB 2; Length 212848;
 Best Local Similarity 91.7%; Pred. No. 5.4e-64;
 Matches 429; Conservative 0; Mismatches 21; Indels 18; Gaps

QY	336 AGCCAGCCCCGACGGCCGATCGGGTCCCTGTCAACCCCGTGGCCACGCCGGCGCCGGAG	395
Db	93875 AGCCAGCCCCGACGGCCGGTCGGTTCCCTGTCAACCCCGTGGCCACGCCGGCGCCGGAG	93934
QY	396 ATCCCCGGATCTGGCAGACCGTAGGCCCGTCTCGTCCGTGACCTTCTGTGGCCTCTC	455
Db	93935 ATCCCCGGATCTGGCAGACCGTAGGCCCGTCTCGTCCGTGA-CTTCTGTGGCCT--	93990
QY	456 CTCCTCACTGGAGGTTGGGGAGGCAGACACCCACCGAACGAGGGAGGGAGCCGG	515
Db	93991 CTCCCTCGTGGAGGGTGGCGG-GGCAGGCAGACACCCAGGAAGGGAGAGGGAGGCC	94049
QY	516 ATCCTCGGGGACCCGGAACCGGAGGAGAGGGTGGCGGAGGAAAGGGTCCCCA	575
Db	94050 ATCCTCGGAGACCCGGGAACCGGAGGCCGAGAGAGGGTGGCGCTGAGGCTGAGGAGGCT	94109
QY	576 GCGCGGAAGCGAGGGAGGGCGATGTTCAAGGCTGAGGCCAGGCCGGGGAGCAGGCC	635
Db	94110 GCCCGGAAGCGAGGGAGGGCGACTGAGGCCGAGAGAGGGTGGCGAGCAGGCC	94169
QY	636 ACCACGGAGGCCAACAGTGGGGGATGCAGCTTGAGCCTGGGAGGAGGCC	695
Db	94170 ACCACGGAGGCCAACAGTGGGGATGCAGCTTGAGCCTGGGAGGAGGCC	94229

RESULT 8
 AC108848/c
 LOCUS AC108848
 DEFINITION Mus musculus clone RP23-300K5, WORKING DRAFT SEQUENCE, 36 ordered pieces.
 ACCESSION AC108848
 VERSION AC108848.2 GI:20336129
 KEYWORDS HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
 SOURCE ORGANISM Mus musculus.
 REFERENCE 1 (bases 1 to 212848)
 AUTHORS Birren,B., Linton,L., Nusbaum,C., LaRocque,K., Lamazares,R.,
 TITLE Mus musculus, clone RP23-300K5
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 212848)
 AUTHORS Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
 Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
 Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
 Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
 Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
 Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
 Landers,T., Lehoczky,J., Levine,R., Liu,G., MacLean,C.,
 McDonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
 McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihowa,T.,
 Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicoll,R., Norbu,C.,
 Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
 Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
 Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
 Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
 Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembekli,L., Zimmer,A. and Zody,M.
 TITLE Direct Submission
 JOURNAL Submitted (31-JAN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 REFERENCE 3 (bases 1 to 212848)
 AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
 Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
 Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
 Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
 Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
 Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G.,
 MacLean,C., McDonald,P., Major,J., Marquis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
 Mihowa,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicoll,R.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
 Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembekli,L., Zimmer,A. and Zody,M.
 TITLE Direct Submission
 JOURNAL Submitted (28-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 COMMENT On Apr 28, 2002 this sequence version replaced g1:18450109. All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 Center project name: L21014
 Center clone name: 300_K5
 ----- Summary Statistics
 Sequencing vector: Plasmid: n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 200069 bases at least 040
 Consensus quality: 206302 bases at least 030
 Consensus quality: 208218 bases at least Q20
 Insert size: 209348; sum-of-contigs
 Quality coverage: 5.7 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 36 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submittor.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 *-----
 * 1 641: contig of 641 bp in length
 * 642 741: gap of 100 bp
 * 742 1427: contig of 686 bp in length
 * 1428 1527: gap of 100 bp
 * 1528 3170: contig of 1643 bp in length
 * 3171 3270: gap of 100 bp
 * 3271 4396: contig of 1126 bp in length
 * 4397 4496: gap of 100 bp
 * 4497 5909: contig of 1413 bp in length
 * 5910 6009: gap of 100 bp
 * 6010 7583: contig of 1574 bp in length
 * 7584 7683: gap of 100 bp
 * 7684 8975: contig of 1292 bp in length
 * 8976 9075: gap of 100 bp
 * 9076 10257: contig of 1182 bp in length
 * 10258 10357: gap of 100 bp
 * 10358 11608: contig of 1251 bp in length
 * 11609 11708: gap of 100 bp
 * 11709 13289: contig of 1581 bp in length
 * 13290 13389: gap of 100 bp
 * 13390 15335: contig of 1946 bp in length
 * 15336 15435: gap of 100 bp
 * 15436 16250: contig of 815 bp in length
 * 16251 16350: gap of 100 bp
 * 16351 17669: contig of 1319 bp in length
 * 17670 17769: gap of 100 bp
 * 17770 19907: contig of 2138 bp in length
 * 19908 20007: gap of 100 bp
 * 20008 21470: contig of 1463 bp in length
 * 21471 21570: gap of 100 bp
 * 21571 23086: contig of 1516 bp in length
 * 23087 23186: gap of 100 bp
 * 23187 24723: contig of 1537 bp in length
 * 24724 24823: gap of 100 bp
 * 24824 27746: contig of 2923 bp in length
 * 27747 27846: gap of 100 bp
 * 27847 29873: contig of 2027 bp in length
 * 29874 29973: gap of 100 bp

Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
 Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G.,
 Scherer, S., Scott, G., Shen, H., Shooshtari, N., Sisson, I.,
 Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
 Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
 Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstock, G. and Gibbs, R.
 Direct Submission
 JOURNAL
 Unpublished
 REFERENCE
 2 (bases 1 to 89551)
 AUTHORS
 Worley, K.C.
 TITLE
 Direct Submission
 COMMENT
 Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 89551)
 Worley, K.C.
 Direct Submission
 Submitted (31-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 Center project name: KAMV
 Center clone name: CH230-69F8
 ----- Summary Statistics
 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 42669 bases at least Q40
 Consensus quality: 46367 bases at least Q30
 Consensus quality: 48722 bases at least Q20
 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 48 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 1151: contig of 1151 bp in length
 1152 1251: gap of unknown length
 1252 2322: contig of 1071 bp in length
 2323 2422: gap of unknown length
 2423 3692: contig of 1270 bp in length
 3693 3792: gap of unknown length
 3793 5136: contig of 1344 bp in length
 5137 5236: gap of unknown length
 5237 6325: contig of 1089 bp in length
 6326 6425: gap of unknown length
 6426 8049: contig of 1624 bp in length
 8050 8149: gap of unknown length
 8150 9419: contig of 1270 bp in length
 9420 9519: gap of unknown length
 9520 10641: contig of 1122 bp in length
 10642 10741: gap of unknown length
 10742 11933: contig of 1192 bp in length
 11934 12033: gap of unknown length
 12034 13362: contig of 1329 bp in length
 13363 13462: gap of unknown length
 13463 14779: contig of 1317 bp in length
 14780 14879: gap of unknown length
 14880 16451: contig of 1572 bp in length
 16452 16551: gap of unknown length
 16552 17915: contig of 1364 bp in length
 17916 18015: gap of unknown length
 18016 19443: contig of 1428 bp in length
 19444 19544: gap of unknown length
 19544 20626: contig of 1083 bp in length
 20627 20726: gap of unknown length
 20727 22100: contig of 1374 bp in length
 22101 22200: gap of unknown length
 22201 23252: contig of 1052 bp in length
 23253 23353: gap of unknown length
 23353 24945: contig of 1593 bp in length
 24946 25045: gap of unknown length
 25046 26886: contig of 1841 bp in length
 26887 26986: gap of unknown length
 26987 28215: contig of 1229 bp in length
 28216 28315: gap of unknown length
 28316 29904: contig of 1589 bp in length
 29905 30004: gap of unknown length
 30005 31803: contig of 1799 bp in length
 31804 31903: gap of unknown length
 31904 33939: contig of 2035 bp in length
 33939 34039: gap of unknown length
 34039 35079: contig of 1040 bp in length
 35079 35179: gap of unknown length
 35179 37353: contig of 2174 bp in length
 37353 37453: gap of unknown length
 37453 39727: contig of 1729 bp in length
 39727 41555: gap of unknown length
 41555 41655: contig of 1616 bp in length
 41655 43271: contig of 2274 bp in length
 43272 43372: gap of unknown length
 43372 44785: contig of 1414 bp in length
 44786 44885: gap of unknown length
 44886 46567: contig of 1682 bp in length
 46567 46668: gap of unknown length
 46668 47949: contig of 2782 bp in length
 47949 49450: gap of unknown length
 49450 49550: contig of 2234 bp in length
 49550 51784: gap of unknown length
 51784 51884: contig of 2163 bp in length
 51884 54046: gap of unknown length
 54047 54146: contig of 2036 bp in length
 54147 56182: gap of unknown length
 56182 56282: contig of 1652 bp in length
 56282 57934: contig of 1652 bp in length
 57934 58034: gap of unknown length
 58034 59825: contig of 1791 bp in length
 59825 59925: gap of unknown length
 59925 61535: contig of 1610 bp in length
 61535 61635: gap of unknown length
 61636 63473: contig of 1838 bp in length
 63474 63573: gap of unknown length
 63574 66376: contig of 2803 bp in length
 66376 66477: gap of unknown length
 66477 68642: contig of 2165 bp in length
 68642 68741: gap of unknown length
 68741 70180: contig of 1439 bp in length
 70180 70280: gap of unknown length
 70280 72432: contig of 2152 bp in length
 72432 72532: gap of unknown length
 72532 73917: contig of 1385 bp in length
 73917 74017: gap of unknown length
 74017 77137: contig of 3120 bp in length
 77137 77237: gap of unknown length
 77237 79997: contig of 2760 bp in length
 79997 80097: gap of unknown length
 80097 82859: contig of 2762 bp in length
 82859 82959: gap of unknown length
 82959 85680: contig of 2721 bp in length
 85680 85780: gap of unknown length
 85780 8581: contig of 3771 bp in length

FEATURES Location/**Qualifiers**
source 1. . 89551

Center: Washington University Genome Sequencing Center
Center code: WUGSC

```

</organism="Rattus norvegicus"
</db_xref="taxon:10116"
</clone="CH230-69F8"

```

Query Match 15.3%; Score 194.8; DB 2; Length 89551;
 Best Local Similarity 82.7%; Pred. No. 6e-30;
 Matches 248; Conservative 0; Mismatches 47; Indels 5; Gaps 2

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

Db	8757	TTTCAGAGTTGAAAGGACTAGATGTTGCCCTGCCGGTCAGACTTCGCCACGTGGACCT	8816
QY	1025	AAACGGCCCCATCGGCAAGACT'TGTGTGGAGATGCAAGGACAACCGCCCTGTCCTGCGAC	1084
Db	8817	AAACGGCCCCATCGTCAAAGACTTGTTGGGAGATGCAAGGACAACGCCCTGTCCTGTGAC	8876

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

QY	1085	AGCACCTTCAGCTTCAAATACATTAGTGAGAGTCGAAACGTTCTGCTA---GA	1140
Db	8877	AGCACCTTCAGCTTCAAATATCATTAGAGAGGTAAAATGGTTCTGCTAAATGGA	8936
Ov	1141	TGGGCTTAATGGAATGACAAGTGAGCTTCTCCCTCTCACCTCTCCCTTCCAAT	1200

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this

QY	1201	TCTTCATGACAGACAGTGTACTTGGATAAAGCCTGTGAATAAAGGTATTGCAAACA	1260
Db	8937	TCGGACAAAGTGGAGCCTCTCCCCACGCCCTCTCCACCTCTCCCTCTCAAAT	8996

sequence, see <http://genome.wustl.edu/gsc>

RESULT 10
AC007970/c
LOCUS AC007970
- - -
138890 bp DNA linear PRI 30-SEP-2000

approach for construction of bacterial artificial chromosomes. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

DEFINITION Homo sapiens BAC clone RP11-485G2 from 2, complete sequence.
ACCESSION AC007970
VERSION AC007970.3 GI:5836168
KEYWORDS HTG.
SOURCE Homo sapiens.

VECTOR: pBACE3:8
NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-349G4. Actual end of this
clone is at base position 138890 of RP11-485G2.
Location/Qualifiers
FEATURES

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 138890)
AUTHORS Sulston, J.E. and Waterston, R.

```
1. .138890
<organism="Homo sapiens"
<db_xref="taxon:9606"
<chromosome="2"
</map="2"
```

TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 138890)

```
/clone="RP11-485G2"  
/clone_lib="RPCI-11"  
663..807  
/rpt_family="L1"  
800..1474
```

AUTHORS Ozersky, P., Kalicki, J. and Sapetti, L.
TITLE The sequence of homo sapiens BAC clone RP11-485G2
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 138890)
ATTACHES Waterston, R. H.

```
/rpt_family="LL"  
1487. .1659  
/rpt_family="LL"  
1798. .2437  
/rpt_family="L1"
```

AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (03-JUL-1999) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis MO 63108, USA

region
2438. .2691
/rpt_family="LL"
2692. .2831
/rpt_family="LL"
3327 3446

REFERENCE 4 (bases 1 to 138890)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (08-SEP-1999) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis

latitude
3327. .3440
/note="match to EST AL046241 (NID:95434325)"
region
4004. .4204
/rpt_family="L2"
4388. .4692
/note="match to EST AL046241 (NID:95434325)"
region

MO 63108, USA
REFERENCE 5 (bases 1 to 138890)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (30-SEP-2000) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT On Sep 8, 1999 this sequence version replaced gi:5630058.
----- Genome Center

repeat_region	/rpt_family="MALK- 6180. .6659
repeat_region	/rpt_family="L2" 6706. .9844
repeat_region	/rpt_family="L1" 9852. .10258
repeat_region	/rpt_family="(TA)n" 10278. .10790

```

repeat_region /rpt_family="L1"
repeat_region /rpt_family="L2"
repeat_region 11461. .17576
repeat_region /rpt_family="L1"
repeat_region 17680. .17758
repeat_region /rpt_family="Mariner"
repeat_region 18412. .18629
repeat_region /rpt_family="MIR"
repeat_region 19942. .20121
repeat_region /rpt_family="L1"
repeat_region 20456. .20629
repeat_region /rpt_family="(TA)n"
repeat_region /rpt_family="L1"
repeat_region 20758. .21114
repeat_region /rpt_family="MALR"
repeat_region 22734. .23101
repeat_region /rpt_family="MALR"
repeat_region 23102. .24334
repeat_region /rpt_family="L1"
repeat_region 24340. .24720
repeat_region /rpt_family="L1"
repeat_region 24721. .25077
repeat_region /rpt_family="L1"
repeat_region 25078. .25115
repeat_region /rpt_family="(TA)n"
repeat_region /rpt_family="L1"
repeat_region 25116. .25599
repeat_region /rpt_family="L1"
repeat_region 25600. .25625
repeat_region /rpt_family="AT-rich"
repeat_region 26058. .26331
repeat_region /rpt_family="L2"
repeat_region 26345. .26475
repeat_region /rpt_family="L2"
repeat_region 27007. .27085
repeat_region /rpt_family="GA-rich"
repeat_region 27113. .27146
repeat_region /rpt_family="(CA)n"
repeat_region 27479. .27605
repeat_region /rpt_family="MER1-type"
repeat_region 27924. .28079
repeat_region /rpt_family="MER1-type"
repeat_region 29818. .30292
repeat_region /rpt_family="MER1-type"
repeat_region 31387. .31688
repeat_region /rpt_family="Alu"
repeat_region 32916. .32981
repeat_region /rpt_family="CT-rich"
repeat_region 32982. .33240
repeat_region /rpt_family="Alu"
repeat_region 33241. .33252
repeat_region /rpt_family="CT-rich"
repeat_region 34110. .34142
repeat_region /rpt_family="AT-rich"
repeat_region 35992. .36014
repeat_region /rpt_family="(TAAA)n"
repeat_region 36034. .36187
repeat_region /rpt_family="Retroviral"
repeat_region 36188. .36481
repeat_region /rpt_family="Alu"
repeat_region 36482. .36514
repeat_region /rpt_family="Retroviral"
repeat_region 36617. .36835
repeat_region /rpt_family="L1"
repeat_region 36838. .36905
repeat_region /rpt_family="(TG)n"
repeat_region 36905. .36939
repeat_region /rpt_family="(TA)n"
repeat_region 36945. .37725
repeat_region /rpt_family="L1"
repeat_region 37736. .38050
repeat_region /rpt_family="Alu"

repeat_region 38070. .38557
repeat_region /rpt_family="MER4-group"
repeat_region 38558. .38976
repeat_region /rpt_family="L2"
repeat_region 39061. .39116
repeat_region /rpt_family="(TAA)n"
repeat_region 39477. .39692
repeat_region /rpt_family="MIR"
repeat_region 40301. .40339
repeat_region /rpt_family="(TC)n"
repeat_region 40405. .40690
repeat_region /rpt_family="Alu"
repeat_region 40713. .40816
repeat_region /rpt_family="MER1-type"
repeat_region 41875. .42095

/note="match to EST AL046241 (NID:g5434325)"
```

Query	Match	Score	DB	Length
Qy	Best Local Similarity	11.6%	147.6	138890
Matches	202; Conservative	73.7%	Pred. No. 3.2e-20;	
Qy	951 TAAGGTGTTACTTCAAACAGTCTGCCGGAGAGGCCCTATGTGTGGGTGTGAGAATCCCTACACCCCTACAGAG	950	890	890
Db	28597 CTGCAAGGACTGTGACATCCATGGGGGTGCTTATGTGTGGGTGTATAGAGACCTAA	28598	28597	28597
Qy	25600. .25625	28541	25600	25625
Qy	/rpt_family="AT-rich"	950	891	891
Db	26058. .26331	890	26058	26331
Qy	/rpt_family="L2"	890	26345. .26475	26345
Db	28597 TAAGAACTACTTCAA---AGTTTACAGAGACTGTGAGACTCTATAACTCGTACCAA	28598	28597	28597
Qy	27007. .27085	28481	27007	27085
Qy	/rpt_family="GA-rich"	28480	27113. .27146	27113
Db	28480 TCTACATAGACCAAATCAGGTTGTAATAACTAGATGTGCTGGCTAGAATTGTCAGAGATGCAAAGAAC	28481	27113	27146
Qy	27479. .27605	28420	27479	27605
Db	/rpt_family="MER1-type?"	28387	27924. .28079	27924
RESULT	11			
AC037443/C				
LOCUS	AC037443			
DEFINITION	Homo sapiens chromosome 2 clone RPL1-77F17 map 2, WORKING DRAFT			
ACCESSION	31387. .31688			
VERSION	32916. .32981			
KEYWORDS	/rpt_family="Alu"			
SOURCE	32982. .33240			
ORGANISM	33241. .33252			
REFERENCE	/rpt_family="CT-rich"			
AUTHORS	34110. .34142			
TITLE	/rpt_family="AT-rich"			
JOURNAL	35992. .36014			
UNPUBLISHED	36034. .36187			
REFERENCE	2 (bases 1 to 151961)			
AUTHORS	Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,			
REFERENCE	Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F., Boguslavskiy,L., Boukhalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G., Locke,K., McDonald,P., Marquis,N., McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheevers,R., Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,			

O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N., Pisan,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE direct Submission

JOURNAL Submitted (08-APR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE 3 (bases 1 to 151961)

AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F., Boguslavkiy,L., Boukhalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Tieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Minova,T., Miranda,C., Mlenga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE Direct Submission

JOURNAL Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT On May 21, 2000 this sequence version replaced gi:7528172.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>

----- project Information

Center project name: L9226
 Center clone name: 77_F_17

----- Summary Statistics

Sequencing vector: M13; M78115; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731

Consensus quality: 147226 bases at least Q40
 Consensus quality: 149421 bases at least Q30
 Consensus quality: 150378 bases at least Q20
 Insert size: 194000; agarose-fp
 Insert size: 151061; sum-of-contigs
 Quality coverage: 4.5 in Q20 bases; agarose-fp
 Quality coverage: 5.8 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 10 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 1451: contig of 1451 bp in length

* 1452 1551: gap of 100 bp

* 1552 3695: contig of 2144 bp in length

* 3696 3795: gap of 100 bp

* 3796 6752: contig of 2957 bp in length

* 6753 6852: gap of 100 bp

Query Match Score 147.6; DB 2; Length 151961;

Best Local Similarity	Pred. No.	DB	Length
73.7%	3.2e-20	202	151961
Conservative	0	Mismatches	69
		Indels	3
		Gaps	1

Matches

QY 831 CTGCAAGACTGCAAATCCGGGGAGAGCGCTATGTGGTGTGCAGGGCACCAG 890

Db 41665 CTGCAAGACTGTGACATCCAATGGGAGGGTGCCTATGTGGTGTATAGAGACCTAA 41606

Db 41605 TAAGAACCTCTCAA--AGTTTACAGAAGCTTGTCAGAAGTGTATACTCGTACCAAG 41549

QY 951 TGGAGGACATCACCTGCAAGTGTAAAGAAACTAGATGGCCTGCCAGTCAGATTIC 1010

Db 41548 GGGAAAGATAACAGCAATCAAAGTGTAAACTAGATGGTGTGGAGATGCAAGGACAAC 41489

QY 1011 GCCACGTGGACCCTAACGCCCATGGGAAGACTTGTGGAGATGCAAGGACAAC 1070

Db 41488 TCTACATAGACCTAAACGGCCCTATGTCAAGATTTGTGGAGATGCAAAGGCAAAC 41429

QY 1071 GCCTGTCTGCACAGCACCTTCAGCTTCAAATA 1104

Db 41428 ACCTATCTGTGAGCACTTCAGCTTCAAATA 41395

REFERENCE 2 (bases 1 to 187578) AC096952. Data from AC032000 and AC024191 were used to finish this
 AUTHORS Wang,C., Meyer,R. and Dignan,G.
 TITLE The sequence of Homo sapiens BAC clone RP11-191J2
 JOURNAL Unpublished (2001)
 REFERENCE 3 (bases 1 to 187578)
 AUTHORS Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (03-OCT-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 4 (bases 1 to 187578)
 AUTHORS Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (29-MAR-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 5 (bases 1 to 187578)
 AUTHORS Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (05-APR-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 6 (bases 1 to 187578)
 AUTHORS Waterston,R.
 TITLE Direct Submission
 JOURNAL Submitted (20-APR-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 COMMENT On Apr 5, 2002 this sequence version replaced gi:19807902.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: sapiens@watson.wustl.edu
 ----- Summary Statistics
 ----- Center project name: H_NH0191J02

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION: Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION: The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E., Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION: The clone sequenced to the left is RP11-473N2; the clone sequenced to the right is RP11-757K22, 2000 bp overlap. Actual start of this clone is at base position 1 of RP11-191J2; actual end is at base position 17325 of RP11-757K22.

Polymorphisms have been identified between AC032000, AC024191, and

```

AC096952. Data from AC032000 and AC024191 were used to finish this
clone, AC096952.
FEATURES source
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /map="4"
  /clone="RP11-191J2"
  /clone_lib="RPCI-11"
  misc_feature
    1..151
    /note="match to EST R34463 (NID:9791364) yh85c01.r1"
  659..698
    /rpt_family="L2"
    1017..1100
    misc_feature
    rpt_family="L1"
    4161..4181
    /note="match to EST BI496001 (NID:g15335345)"
  4163..4477
    misc_feature
    rpt_family="Alu"
    4178..4188
    /note="match to EST AA749271 (NID:g2789229) ny04d05.s1"
  4541..4567
    misc_feature
    rpt_family="(TTTG)n"
    4557..4577
    /note="similar to Homo sapiens EST AI265985 (NID:g3884143)
  qm46c01.x1"
  5981..5992
    /note="match to EST BI493951 (NID:g15333295)"
  8924..9119
    rpt_family="L1"
  9141..10080
    rpt_family="L1"
    10292..10618
    rpt_family="L1"
    10536..11223
    /note="match to EST BG506934 (NID:g13468451)"
  10662..10970
    rpt_family="L1"
  11491..11819
    rpt_family="L1"
  11832..12135
    rpt_family="Alu"
  12137..12205
    rpt_family="MER2_type"
  12204..12819
    rpt_family="MER2_type"
  12830..13102
    rpt_family="L1"
  13109..13255
    rpt_family="MER2_type"
  13256..13378
    rpt_family="L1"
  13379..13664
    rpt_family="Alu"
  13665..14289
    rpt_family="L1"
  14301..14515
    /note="match to EST BG995513 (NID:g14399583)"
  14357..14790
    /note="similar to Homo sapiens EST BE735137
  (NID:g10149129)"
  14405..14880
    /note="similar to Mus musculus EST BI407638
  (NID:g15168561)"
  14409..14845
    /note="similar to EST BM912100 (NID:g19362479)"
  14441..14845
    /note="similar to EST BQ054284 (NID:g19813624)"
  14447..14827
    /note="similar to Homo sapiens EST BE799707
  (NID:g10220905)"
  14447..14813
  misc_feature
```

	SOURCE ORGANISM	Homo sapiens.
misc_feature	/note="similar to Homo sapiens EST BE378224 (NID:g9323589)"	1447..14761
misc_feature	/note="similar to Homo sapiens EST BF695420 (NID:g11980828)"	14464..14827
misc_feature	/note="similar to Homo sapiens EST AT815933 (NID:g5431479) au43f12.x1"	14467..14880
misc_feature	/note="similar to Mus musculus EST BE848577 (NID:g10306847) uw38h07.y1"	14477..14827
misc_feature	/note="similar to Homo sapiens EST AA143121 (NID:g1712509) z149a07.rl"	14544
misc_feature	/note="match to EST BG995513 (NID:g14399583)"	14623..14884
misc_feature	/note="similar to Homo sapiens EST BG034944 (NID:g12428670)"	14623..14845
misc_feature	/note="similar to EST BQ021037 (NID:g19756307)"	14623..14827
misc_feature	/note="similar to Homo sapiens EST AW675060 (NID:g7540295) ba60h07.y1"	14671..14827
repeat_region	/rpt_family="CT-rich" 14733..14769	14858..15138
misc_feature	/note="similar to EST BM740442 (NID:g19061771)"	14873..14900
repeat_region	/rpt_family="(TG)n"	14878..15152
misc_feature	/note="similar to EST BM912100 (NID:g19362479)"	14878..15146
misc_feature	/note="similar to EST BQ021037 (NID:g19756307)"	14878..15146
misc_feature	/note="similar to EST BQ054284 (NID:g19813624)"	10..5%; Score 134..6; DB 9; Length 694; Best Local Similarity 68..7%; Pred. No. 7..1e-05; Matches 138; Conservative 0; Mismatches 44; Indels 19; Gaps 2;
Qy	140 TCCTCCCCGGCTACAGACAGCTCATGGCCGGAGTACGTCGACAGGCCACCCAGGGCAC 199	Db 528 TCTTTCCGGGTGGGGCGTCTGACGGCCGGAGTA-TTTGACAGCTACCCAGGGAGC 470
Qy	200 AGCTCATGGCCCTGCTGTCGGGATGGGT-----CCGGTGGTCA 241	Db 469 GGCTCATGGCTCTGGCGAGGTGGGGGGTCTGGACCCGGGGGGCAGGTCCG 410
Qy	242 GCAGCCGGTGCCTGGTGCAGGTGAACCCGGCCGACGCCCTGGTGCAGTGTCA 301	Db 409 GCAGCCGGTGCCTGGTGCAGGTGAACCCGGCCGACGCCCTGGTGCAGTGTCA 350
Qy	302 TCGGGGCCACGGTGCAGGTGACGCCGTCAGCAGCCGGTACAGTGTCCG 322	Db 349 TGGGGAGGCCACGGTGCAGC 329
RESULT 15	AF429315/c	
LOCUS	AF429315	125020 bp
DEFINITION	Homo sapiens junctophilin 3 (JPH3)	DNA linear
ACCESSION	AF429315	gene, partial cds.
VERSION	AF429315.1	GI:17646244
KEYWORDS		
SOURCE	Homo sapiens.	
ORGANISM	Homo sapiens.	
REFERENCE	Homo sapiens; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 125020)	
AUTHORS	Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S., Ingersoll-Ashworth,R.G., Fleisher,R., Stevanin,G., Brice,A., Potter,N.T., Ross,C.A. and Margolis,R.L.	
TITLE	A repeat expansion in the gene encoding junctophilin-3 is associated with Huntington disease-like 2	
JOURNAL	Nat. Genet. 29 (4), 377-378 (2001)	
MEDLINE	21583737	
PUBMED	11694876	
REFERENCE	2 (bases 1 to 125020)	
AUTHORS	Holmes,S.E., Ingersoll-Ashworth,R.G., Ross,C.A. and Margolis,R.L.	
TITLE	Direct Submission	
JOURNAL	Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical	
RESULT 14		
HSA335296/c	HSA335296	694 bp DNA linear PRI 18-JUL-2002
LOCUS	Homo sapiens genomic sequence surrounding NotI site, clone	
DEFINITION	NR5-ID16C.	
ACCESSION	AJ335296	
VERSION	AJ335296.1	GI:15879714
KEYWORDS		

FEATURES SOURCE	Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA Location/Qualifiers	Db 17243 NNRMGKTCNCNYMRRSRAMMNAAAAGCTCCCCANTNGGGAAAGGGCGSASRAS 17184
1. .125020	/organism="Homo sapiens" /db_xref="taxon: 9606"	QY 600 TCAGGCTGCAGGGCAGGCCGGGGAGCAGCCACCGGAGGACCGAACAGCTGT 659
/chromosome="16"	/map="16q24.3; between D16S520 and WI-12410" /note="isolated from a patient with Huntington's disease-Like 2 (HDL2)"	Db 17183 CYKGRMSSKSCYRSRSGTRRRCMKSKRRCGSRGKSMGTRGRSGGKTSYSAKGSGRGCVW 17124
repeat_region	complement(35581. .35746)	/rpt_type=tandem
gene	/rpt_unit=ctg complement(<36507. .>36887)	/gene="JPH3"
mRNA	/note="JPH3"	/product="junctionophilin 3"
CDS	complement(<36507. .36887)	/gene="JPH3"
	/note="component of the junctional complex between plasma membrane and endoplasmic reticulum" /codon_start=1	/product="junctionophilin 3"
	/protein_id="AAL40941.1" /db_xref="GI:17646245"	/protein_id="AAL40941.1"
	/translation="MSSGGRNFDDGGSYCGWEDGKAHGHGVCTGPKGQGEYTGSWSHGFEVLGVITWPMSGNTQGTWAQGKRHIGLESKGKWVYKGEWTHGFKGKGRYGVRECAGNGAKVEGTWSNGLQDGYGTETYSQG"	
BASE COUNT	29056 a 32731 c 30696 g 28283 t 4254 others	
ORIGIN		
Query Match	5.7%; score 72.2; DB 9; Length 125020;	
Best Local Similarity	11.1%; Pred. No. 0.00011; Mismatches 395; Indels 4; Gaps 3;	
Matches	100; Conservative 401; Mismatches 395; Indels 4; Gaps 3;	
QY 4	GGGGGGAGGGCGGGACGCACCCATGTTCCCGGAGCACGTTCCACCCCTGCCGCAT 63	QY 720 GGACCCCGGTGATTGGATGCCCTCGAGAACCCAGGCCTCCCCGCAAAGCACGGAGCAGGA 779
Db 17783	GSKSSRGGTGYYMKKKGGKGMSSTSRRRGSSAKSCSYMWMGMSMCMSMA 17724	Db 17063 MGMSSYSSKGYSWGMKSYMWRSYYSKRSTSKAWRSSLRKGMTGGRSGYKGKGGGS 17004
QY 64	CCTTATCCGGAGGCCACCAAGCGGGATGGCTGGAGGTTCCAGGCCAGGGCTGCCGA 123	QY 780 CAAGGAGGCCAGGGCTTCCAGTTAGACAGAACAGTACGGCTACTACTGCAAGGA 839
Db 17723	KSYMMCYCIMYRMSYMSYYKCYSCMGMSSTSYSYCCWKMSSWGSCYKCMKYSGWSS 17664	Db 17003 YKGGGSWGGKGKSTRGSSAGKKSYKMSCCARYKSMSYKCYSSRWMCMSYYCWGGK 16944
QY 124	CCCGGCCCCCTCCTCCCCGGCTACAGACAGCTCATGGCCGGAGTACGTCA- 182	QY 840 CTGCAAATCCGGTGGAGAGGCCATATGTGGTGTGCAAGGGCACCAAGTAAGGTGTT 899
Db 17663	YSTMGMYYSTCKYKCSWSSMYKCKTSKSYRKRSYYWGKRAKKYYCAGRRRRM 17604	Db 16943 CWSSCWWSSMRKSKGGSWAKGMWDKGSVSTDKSDKMBRSBSKVKSKWSMSRRGKYY 16884
QY 183	-CAGCCACCAGGGCACAGCTCATGGCCCTGCTGCGGGATGGGTCCCCGGTCGGCA- 241	
Db 17603	SYWKCCAkwmsyccwsyscmttyskstckykrkgyywggsktcysaggksrsmyy 17544	
QY 242	GCAGCCGTGACGCGTGCAG-GTGAACCCGCCGACGCCCTCGGTGCAGTGTCA 300	
Db 17543	CMMRSSKSSWMSMARSSWCMGNGRKSAGWGRSSKGKRSKTMKRAKSKKT 17484	
QY 301	CTCGGGGCCGCGCACGCTGCAGGCTGCAGGGTGCCAGGCCAGCCCCGACGCCGATCGGGT 360	
Db 17483	GSYGSTGRSMKKKGSKYKSRGMKGKKTCYCMWKKYKRTSMCWYIMKSWGYKRYR 17424	
QY 361	TCCTGTCAACCCCGTGGCCACGCCGGGGAGATCCCCCG-AATCCITGGCAGACCGT 419	
Db 17423	CCMKKKGCTGYRGMSKKSKYKSRGMGRGSSYSTSCWKSCWGYSWMKCMKMSYKKRR 17364	
QY 420	AGCCCCGTTCTGTCGGTGAACCTCTGTGGCCTCTCTCTACTGGAGGTGCGGGAGG 479	
Db 17363	MRGSMSKKGWRYGAGRGRGYCYSWMSMSTRRKSCKYSYKSKRGSKGRSTKSKAKSSMRMAGSKCTGYSSYWSNNR 17304	
QY 480	CAGGAGACACCCACGAAGGGAGGGAGGCCATCTCGGGACCCGGAAACCGGA 539	
Db 17303	SMKKMRRSSWCKYKSRGSRRKSKGWGRSTKSKAKSSMRMAGSKCTGYSSYWSNNR 17244	
QY 540	GCCGAGAGGGTGGCCGAGGAAGCGGTCCCCCAGCCGCGAAGCGAGGGCGATGT 599	

Search completed: January 19, 2003, 22:13:21
Job time : 5277 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on : January 18, 2003, 09:41:06 ; Search time 76 Seconds

(without alignments)

5152.977 Million cell updates/sec

Title: US-09-844-864-1

Perfect score: 1277

Sequence: 1 aaggcgggcggccgggaa.....aaaaaaa.....aaaaaaa 1277

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_patents_NA:*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*

5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*

6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. 1s is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and 1s derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	54	4.2	5467	Sequence 12, Appl
2	54	4.2	5467	Sequence 12, Appl
3	53.2	4.2	7218	Sequence 12, Appl
4	50.8	4.0	6232	Sequence 14, Appl
5	49.6	3.9	400	Sequence 11, Appl
6	49.6	3.9	400	Sequence 179, App
7	49	3.8	715	Sequence 174, App
8	49	3.8	715	Sequence 8, Appl
9	49	3.8	715	Sequence 8, Appl
10	49	3.8	715	Sequence 8, Appl
11	49	3.8	715	Sequence 8, Appl
12	49	3.8	715	Sequence 8, Appl
13	49	3.8	715	Sequence 8, Appl
14	49	3.8	715	Sequence 8, Appl
15	49	3.8	715	Sequence 8, Appl
16	49	3.8	715	Sequence 8, Appl
17	49	3.8	7266	Sequence 8, Appl
18	49	3.8	7262	Sequence 8, Appl
19	49	3.8	7362	Sequence 7, Appl
20	49	3.8	7362	Sequence 7, Appl
21	49	3.8	7362	Sequence 7, Appl
22	49	3.8	7362	Sequence 7, Appl
23	49	3.8	7362	Sequence 7, Appl
24	49	3.8	7362	Sequence 7, Appl
25	49	3.8	7362	Sequence 7, Appl
26	49	3.8	7364	Sequence 7, Appl
27	49	3.8	7376	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-07-745-206A-12
Sequence 12, Application US/07745206A
Patent No. 5429921

GENERAL INFORMATION:

APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: McCue, Ann
APPLICANT: Feldman, Daniel
TITLE OF INVENTION: Human Calcium Channel Compositions and Methods
TITLE OF INVENTION: Methods
NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:
ADDRESSEE: Fitch, Even, Rabin & Flannery
STREET: 135 S. Lasalle
CITY: Chicago
STATE: Illinois
COUNTRY: U.S.A.
ZIP: 60603

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/745,206A
FILING DATE: 19910815
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Feder, Scott B
REFERENCE/DOCKET NUMBER: 51504
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-372-7842
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 5467 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)

FEATURE:
NAME/KEY: CDS
LOCATION: join(144..3164, 3168..3245, 3249..3386, 3390
LOCATION: ..3392..3396..3488, 3495..3539, 3543..3581, 3585
LOCATION: ..3587..3591..3626, 3630..3689, 3693..3737, 3744
LOCATION: ..3746..3750..4823, 4827..4841, 4845..5006, 5010
LOCATION: ..5096..5100..5306, 5310..5366, 5370..5465)
US-07-745-206A-12

Page 2

SEQUENCE CHARACTERISTICS:
 LENGTH: 5467 base pairs
 TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: join(144..3164, 3168..3245, 3249..3386, 3390
LOCATION: ..3392, 3396..3488, 3495..3539, 3543..3581, 3585..
LOCATION: ..3587, 3591..3626, 3630..3689, 3693..3737, 3744..
LOCATION: ..3746, 3750..4823, 4827..4841, 4845..5006, 5010..
LOCATION: ..5096, 5100..5306, 5310..5366, 5370..5465)

Query	Match	Score	DB	Length
Qy	4.2%	54	2	346;
Best Local Similarity	46.7%	Pred. No.	0.00048	
Matches	171;	Conservative	0;	Mismatches 195; Indels 0; Gaps
Qy	303	CGGGGCCGCACGCTGCAGGGTGCAGCCAGCCCCGACGCCGATCGGGTTC	362	
Db	2800			
Qy	363	CTGTCAACCCCGTGGCCACGCCGGGGAGATCCCCGCGATCCTGGCAGACCGTAGC	422	
Db	2860			
Qy	423	CGGCGCCAGGGCCCAGGGCCGAGGGCGGGCACACCGGGCGGGCTCCCCGGAGGA	2919	
Db	2920			
Qy	483	GGCGGCCGAGGGAGCCCCGACGCCACCGCGCGCACGGCACCGAGGATCCGAGCAAGGA	542	
Db				

US-08-311-363-12
Sequence 12, Application US/08311363
Patent No. 5876958

GENERAL INFORMATION:

APPLICANT: Harpold, Michael
APPLICANT: Ellis, Steven
APPLICANT: Williams, Mark
APPLICANT: Feldman, Daniel
APPLICANT: McCue, Ann
APPLICANT: Brenner, Robert

TITLE OF INVENTION: Human Calcium Channel Compositions and Methods

NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:

QY	543	GAGAGAGGTGGCGGAGGAACGGTCCCCAGCCGGAACGGAGGGATGTTCA	602
Db	3040	CGGGAGGGAGAGCGGGAGGCCGGCGGGCACAGGCACAAGGCGCA	3099
QY	603	GGCTGCAGGGCAGGCCGGTGGAGCAGCAGCCACCACCGGAGGACCGGAAACAGTGTGGC	662
Db	3100	GCCTGCTCACGAGGCTGTGGAGAACGGAGACCACGGAGGAAGGGGCCACGGAGAAGGGAGGC	3159
QY	663	GGCGAT 668	
Db	3160	TGAGAT 3165	

RESULT 5
 US-09-056-556-179
 Sequence 179, Application US/09056556
 ; Patent No. 6350456
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, Steven G.
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Dillon, Davin C.
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND
 ; NUMBER OF SEQUENCES: 241
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SEED and BERRY LLP
 ; STREET: 6300 Columbia Center, 701 Fifth Avenue
 ; CITY: Seattle
 ; STATE: Washington
 ; ZIP: 98104-7092
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/056, 556
 ; FILING DATE: 07-APR-1998
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Maki, David J.
 ; REGISTRATION NUMBER: 31, 392
 ; REFERENCE/DOCKET NUMBER: 210121.457
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (206) 622-4900
 ; TELEFAX: (206) 682-6031
 ; INFORMATION FOR SEQ ID NO: 179:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 400 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-09-056-556-179

Query Match 3.9%; Score 49.6; DB 4; Length 400;
 Best Local Similarity 49.2%; Pred. No. 0.0021;
 Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

```

QY 458 CCTCACTGGAGGTGGGGAGGCAGGCCAGACACCCACGAGGGAGGGAGCCGGCAT 517
Db 119 CTTCAACCAAGGGCGCGACGGCAACGGCGCAACGGCGGTGACGGGGGTGGCGCAA 178
QY 518 CCTGGGACCCGGAACCGGGAGAGGGTGGCCGGAGGAAGGGTCCCCCAGC 577
Db 179 CCTGGGAAACGGGGAACCGGGCAGACAACACCACCGCCGGCCACACAGG 238
QY 578 CGCGAAGGGAGGGGATGTCAGGCTGCAGGGCAGGGGGGGGGAGCAGCCAC 637
Db 239 CGGGGACGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGCAC 298
QY 638 CACCGGAGGACGGAACAGTGTGGGGCGATGCACTGGGGAGGGAGGCCAT 697
Db 299 CGGCACGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGCAC 358
QY 698 GTCTGGCCGAGAGATGGCTCAGG 721
Db 359 CGGGGGGACGGGTGCACTCTCAGG 382

```

RESULT 6
 US-09-072-596-174
 Sequence 174, Application US/09072596
 ; Patent No. 6458366
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, Steven G.

RESULT 7
 US-08-455-543A-8
 Sequence 8, Application US/08455543A
 ; Patent No. 5792846
 ; GENERAL INFORMATION:
 ; APPLICANT: Harpold, Michael
 ; APPLICANT: Ellis, Steven

	Matches	201;	Conservative	0;	Mismatches	215;	Indels	3;	Caps	2;
APPLICANT:	Williams, Mark									
APPLICANT:	Feldman, Daniel									
APPLICANT:	McCue, Ann									
APPLICANT:	Brenner, Robert									
TITLE OF INVENTION:	HUMAN CALCIUM CHANNEL COMPOSITIONS AND METHODS									
NUMBER OF SEQUENCES:	57									
CORRESPONDENCE ADDRESS:										
ADDRESSEE:	Brown, Martin, Haller & McClain									
STREET:	1660 Union Street									
CITY:	San Diego									
STATE:	California									
COUNTRY:	USA									
ZIP:	92101-2926									
COMPUTER READABLE FORM:										
MEDIUM TYPE:	Diskette									
COMPUTER:	IBM Compatible									
OPERATING SYSTEM:	DOS									
SOFTWARE:	FastSEQ Version 1.5									
CURRENT APPLICATION DATA:										
APPLICATION NUMBER:	US/08/455,543A									
FILING DATE:	May 31, 1995									
PRIOR APPLICATION DATA:										
APPLICATION NUMBER:	08/223,305									
FILING DATE:	April 4, 1994									
PRIOR APPLICATION DATA:										
APPLICATION NUMBER:	07/868,354									
FILING DATE:	April 10, 1992									
PRIOR APPLICATION DATA:										
APPLICATION NUMBER:	US 07/745,206									
FILING DATE:	15-AUG-1991									
PRIOR APPLICATION DATA:										
APPLICATION NUMBER:	US 07/620,250									
FILING DATE:	30-NOV-1990									
PRIOR APPLICATION DATA:										
APPLICATION NUMBER:	US 07/482,384									
FILING DATE:	20-FEB-1990									
PRIOR APPLICATION DATA:										
APPLICATION NUMBER:	US 07/603,751									
FILING DATE:	04-APR-1989									
PRIOR APPLICATION DATA:										
APPLICATION NUMBER:	WO PCT/US89/01408									
FILING DATE:	04-APR-1989									
PRIOR APPLICATION DATA:										
APPLICATION NUMBER:	US 07/176,899									
FILING DATE:	04-APR-1988									
ATTORNEY/AGENT INFORMATION:										
NAME:	Seidman, Stephanie L.									
REGISTRATION NUMBER:	33,779									
REFERENCE/DOCKET NUMBER:	6362-52517									
TELECOMMUNICATION INFORMATION:										
TELEPHONE:	(619)238-0999									
TELEFAX:	(619)238-0062									
INFORMATION FOR SEQ ID NO:	8:									
SEQUENCE CHARACTERISTICS:										
LENGTH:	7175 base pairs									
TYPE:	nucleic acid									
STRANDEDNESS:	double									
TOPOLOGY:	linear									
MOLECULE TYPE:	DNA (genomic)									
FEATURE:										
NAME/KEY:	CDS									
LOCATION:	144..6857									
FEATURE:										
NAME/KEY:	5'UTR									
LOCATION:	1..143									
FEATURE:										
NAME/KEY:	3'UTR									
LOCATION:	6855..7175									
US-08-455-543A-8										
Query Match	3.8%	Score 49;	DB 1;	Length 7175;						
Best Local Similarity	48.0%	Score 49;	DB 1;	Length 7175;						
Pred. No.	0.0094;									

RESULT 8

US-08-193-078B-8

Sequence 8, Application US/08193078B

Patent No. 5846757

GENERAL INFORMATION:

APPLICANT: Harpold, Michael

APPLICANT: Ellis, Steven

APPLICANT: Williams, Mark

APPLICANT: Feldman, Daniel

APPLICANT: McCue, Ann

APPLICANT: Brenner, Robert

TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND METHODS

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: BROWN, MARTIN, HALLER & MCCLAIN

STREET: 1660 UNION STREET

CITY: SAN DIEGO

STATE: CA

COUNTRY: USA

ZIP: 92101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/193,078B

FILING DATE: 07-FEB-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/868,354

FILING DATE: 10-APR-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/745,206

FILING DATE: 15-AUG-1991

ATTORNEY/AGENT INFORMATION:

NAME: Seidman, Stephanie L.

REGISTRATION NUMBER: 33,779

REFERENCE/DOCKET NUMBER: 6362-53607

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-238-0999

ADDRESSEE: Brown, Martin, Haller & McLain
 STREET: 1660 Union Street
 CITY: San Diego
 STATE: California
 COUNTRY: US
 ZIP: 92101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/949,386
 FILING DATE:
 CLASSIFICATION:

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/290,012
 FILING DATE: 11-AUG-1994
 APPLICATION NUMBER: 08/149,097
 FILING DATE: 5-NOV-1993

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/105,536
 FILING DATE: 11-AUG-1993

ATTORNEY/AGENT INFORMATION:
 NAME: Seidman, Stephanie L.
 REGISTRATION NUMBER: 33,779
 REFERENCE/DOCKET NUMBER: 519808

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (619) 238-0999
 TELEFAX: (619) 238-0062

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:
 LENGTH: 7175 base Pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:
 NAME/KEY: CDS
 LOCATION: 144..6857

FEATURE:
 NAME/KEY: 5'UTR
 LOCATION: 1..143

FEATURE:
 NAME/KEY: 3'UTR
 LOCATION: 6855..7175

US-08-949-386-8

Query Match 3.8%; Score 49; DB 3; Length 7175;
 Best Local Similarity 48.0%; Pred. No. 0.0094; Mismatches 215; Indels 3; Gaps 2;

Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY 250 GACCGCTGCGGTGCAGGTGAAACCCGGCCGAGCCTCGGTGCACTCGGGCGC 309
 Db 2751 GAGGCCCGAAGGGCGAGAGGGGGAGGCCCGGTGCCGGAGGCCGCCAC 2810

QY 310 CGCACCGCTGCAGCCTGCAGGGTGCCGAGCCAGCCCCGATCGGGTTCTGTCAA 369
 Db 2811 CGCA-GCCACAGCAAGGAGGCCGCCGCAC 2867

QY 370 CCCCGTGGCCACGCCGGGGAGATCCCCGGCATCTGGAGCTGGCAGACCGTAGCCCCCGTTC 429
 Db 2868 GGCCCCAGGCCGGAGGGGGAGGCCGCCACCGGGCGGGCTCCCCGGAGGGCGGCC 2927

QY 430 TCGTCCTGACCTCTGGCTCTCCCTCACTGGAGTTGGGAGGCCAGACA 489
 Db 2928 GAGCGGGAGCCCGACGCCACGGCGCACCAGGATCCGAGCAAGGAGTGGCC 2987

QY 490 CCCACGAAGGGAGAGGGGGAGCCGGCATCTGGGACCCGGAAACCGGAGGCCAGAG 549
 Db 2988 GGGCCAAGGGAGGGGGAGGCCGCCGGAGGGGGGGGGGGGGGGGGAGGGGGGGGGAG 3047

QY 550 GTGGCCGGAGGGAAAGCGGTCCCCCAGCCGAAAGCGGAGGGCGATGTCAGGCTGCA 609

RESULT 12
 US-08-450-562-8
 Sequence 8, Application US/08450562
 Patent No. 6096514

GENERAL INFORMATION:

APPLICANT: Harpold, Michael
 APPLICANT: Ellis, Steven
 APPLICANT: Williams, Mark
 APPLICANT: McCue, Ann
 APPLICANT: Gillespie, Alison
 APPLICANT: Feldman, Daniel
 APPLICANT: Brenner, Robert

TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND METHODS

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Brown, Martin, Haller & McLain
 STREET: 1660 Union Street
 CITY: San Diego
 STATE: California
 COUNTRY: US
 ZIP: 92101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/450,562
 FILING DATE:
 CLASSIFICATION: 435

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/404,950
 FILING DATE: 13-MAR-1995
 APPLICATION NUMBER: 08/336,257
 FILING DATE: 7-NOV-1994

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/314,083
 FILING DATE: 28-SEPT-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/311,363
 FILING DATE: 23-SEPT-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/290,012
 FILING DATE: 11-AUG-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/223,305
 FILING DATE: 4-APR-1994

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/193,078
 FILING DATE: 07-FEB-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/149,097
 FILING DATE: 5-NOV-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/105,536
 FILING DATE: 11-AUG-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/914,231
 FILING DATE: 13-JULY-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 07/868,354
 FILING DATE: 10-APR-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US92/06903

Db 3048 GCGGAGAGGGAGCCGGGGGGGACCGGGGGCACAGGGCAGCCCTGCT 3107
 QY 610 GGGCAGGCCGGTGGAGCAGCAGCCACCGGAGGAGAACAGTGTGGGGCAT 668
 Db 3108 CACGAGGCTGTGGAGAGGAGACCGGAGAAGGGAGGGGGCTGAGAT 3166

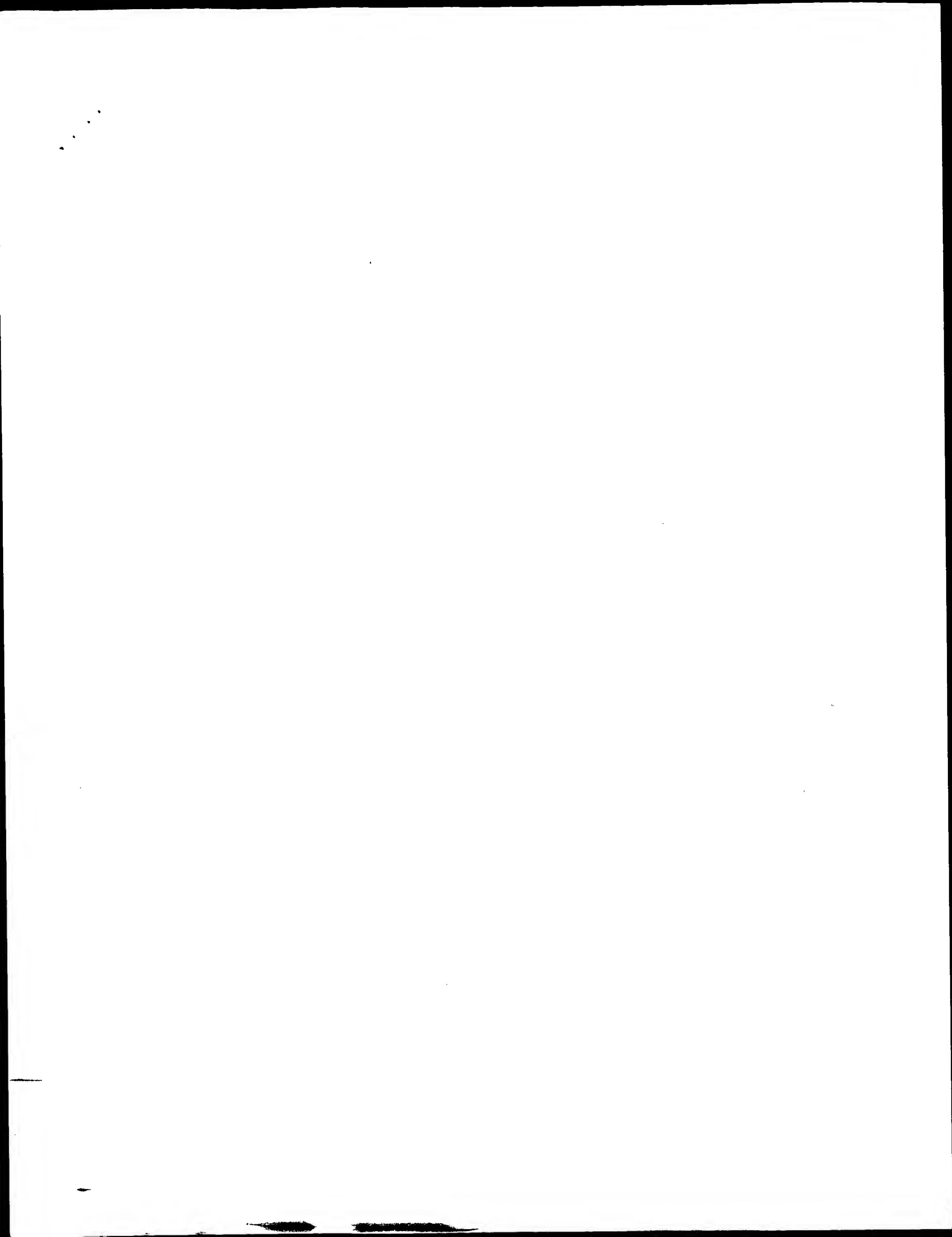
RESULT 15
 US-09-268-163-7
 ; Sequence 7, Application US/09268163B
 ; Patent No. 6353091
 ; GENERAL INFORMATION:
 ; APPLICANT: Lipscombe, Diane
 ; APPLICANT: Schorge, Stephanie
 ; TITLE OF INVENTION: HUMAN N-TYPE CALCIUM CHANNEL ISOFORM AND USES THEREOF
 ; FILE REFERENCE: B105/7000
 ; CURRENT APPLICATION NUMBER: US/09/268,163B
 ; CURRENT FILING DATE: 1999-03-12
 ; EARLIER APPLICATION NUMBER: US 60/077,901
 ; EARLIER FILING DATE: 1998-03-13
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: FASTSEQ for Windows Version 3.0
 ; SEQ ID NO 7
 LENGTH: 7177
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 146..6856
 ; US-09-268-163-7

Query Match 3.8%; Score 49; DB 4; Length 7177;
 Best Local Similarity 48.0%; Ptd. No. 0.0094; Mismatches 2;
 Matches 201; Conservative 0; Indels 3; Gaps 2;

```

Qy  250 GACCTGCGCTGAGGTGAAACCCGGGGGAGCCCTGGTCAATGTTCACTCGGGGGC 309
Db  2753 GAGGCCCGAAGGGGGAGGCGGGAGCCGGTGCCTGGGAGGGAGGGGGGGGGCAC 2812
Qy  310 CGCACGCTGCAGCCTGCAGGGTGCAGCCCCGACGCCGATCGGGTCTGTCAA 369
Db  2813 CGCA-GCCACAGCAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGA 2869
Qy  370 CCCCGTGGCACGGGGGGAGATCCCCGGATCCTGGCAGACCGTAGCCCCGTT 429
Db  2870 GGCCCAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2929
Qy  430 TCGTCCGTGACCTCTGTGGCTCTCTCTCACTGGAGGTGCGGGAGGCAGGCAGACA 489
Db  2930 GAGGGGGAGGCCGACGCCACGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2989
Qy  490 CCCACGAGGGAGGGAGGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 549
Db  2990 GGGGCCAAGGGGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 3049
Qy  550 GTGGCCGGAGGAAGCGGTCCCCCAGCCGCGAAGGGAGGGGAGATGTTCAGGCTGCA 609
Db  3050 GCGGAGAGGGGGAGGCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 3109
Qy  610 GGGCAGGGGGGGGGAGCAGCAGCCACCCGGAGGACCGAACAGTGTGGGGGAT 668
Db  3110 CACGAGCTGTGAGAAGGGAGGCCACGGAGAACGGAGGAGAT 3168

```



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 19, 2003, 19:49:35 ; Search time 62 seconds

(without alignments)

9183.449 Million cell updates/sec

Title: US-09-844-864-1

Perfect score: 1277

Sequence: 1 aaggcgccgaggcgccggaa.....acaaaaaaaaaaaaaa 1277

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 393868 seqs, 222934149 residues

Total number of hits satisfying chosen parameters:

787736

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_NA:*

/cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*

/cgn2_6/ptodata/1/pct_new_pub.seq:*

/cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*

/cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*

/cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*

/cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*

/cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*

/cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*

/cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*

/cgn2_6/ptodata/1/pubpna/us10_NEW_PUB.seq:*

/cgn2_6/ptodata/1/pubpna/us10_PUBCOMB.seq:*

/cgn2_6/ptodata/1/pubpna/us60_NEW_PUB.seq:*

/cgn2_6/ptodata/1/pubpna/us60_PUBCOMB.seq:*

ALIGNMENTS

RESULT 1

US-09-844-864-1
Sequence 1, Application US/09844864
Patent No. US20020042926A1

GENERAL INFORMATION:

APPLICANT: Matzuk, Martin

APPLICANT: Ren, Yongsheng

APPLICANT: Wu, Xuemei

TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS

FILE REFERENCE: P01925US2 / 0980779 / OTA 99-48

CURRENT APPLICATION NUMBER: US/09/844,864

CURRENT FILING DATE: 2001-04-27

PRIORITY FILING DATE: 1998-10-28

PRIORITY APPLICATION NUMBER: 60/106,020

NUMBER OF SEQ ID NOS: 25

SOFTWARE: PatentIn version 3.0

SEQ ID NO 1
LENGTH: 1277
TYPE: DNA
ORGANISM: Mus musculus

US-09-844-864-1

Query Match 100.0% ; Score 1277; DB 10; Length 1277;
Best Local Similarity 100.0%; Pred. No. 2e-307;
Matches 1277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1277	100.0	1277	10	US-09-844-864-1 Sequence 1, Appli Sequence 18, Appli Sequence 22, Appli Sequence 21, Appli Sequence 25, Appli Sequence 19, Appli Sequence 23, Appli Sequence 20, Appli Sequence 24, Appli Sequence 1601, Appli Sequence 7, Appli Sequence 23, Appli Sequence 20, Appli Sequence 24, Appli Sequence 1601, Appli Sequence 7, Appli Sequence 1179, Appli Sequence 5, Appli Sequence 3, Appli Sequence 104, Appli Sequence 104, Appli Sequence 104, Appli Sequence 35, Appli Sequence 1, Appli
2	799.4	52.6	814	10	US-09-844-864-18 Sequence 1, Appli Sequence 18, Appli Sequence 22, Appli Sequence 21, Appli Sequence 25, Appli Sequence 19, Appli Sequence 23, Appli Sequence 20, Appli Sequence 24, Appli Sequence 1601, Appli Sequence 7, Appli Sequence 1179, Appli Sequence 5, Appli Sequence 3, Appli Sequence 104, Appli Sequence 104, Appli Sequence 104, Appli Sequence 35, Appli Sequence 1, Appli
3	752.2	58.9	809	10	US-09-844-864-22 Sequence 1, Appli Sequence 18, Appli Sequence 22, Appli Sequence 21, Appli Sequence 25, Appli Sequence 19, Appli Sequence 23, Appli Sequence 20, Appli Sequence 24, Appli Sequence 1601, Appli Sequence 7, Appli Sequence 1179, Appli Sequence 5, Appli Sequence 3, Appli Sequence 104, Appli Sequence 104, Appli Sequence 104, Appli Sequence 35, Appli Sequence 1, Appli
4	273.6	21.4	305	10	US-09-844-864-21 Sequence 1, Appli Sequence 18, Appli Sequence 22, Appli Sequence 21, Appli Sequence 25, Appli Sequence 19, Appli Sequence 23, Appli Sequence 20, Appli Sequence 24, Appli Sequence 1601, Appli Sequence 7, Appli Sequence 1179, Appli Sequence 5, Appli Sequence 3, Appli Sequence 104, Appli Sequence 104, Appli Sequence 104, Appli Sequence 35, Appli Sequence 1, Appli
5	260.6	20.4	375	10	US-09-844-864-25 Sequence 1, Appli Sequence 18, Appli Sequence 22, Appli Sequence 21, Appli Sequence 25, Appli Sequence 19, Appli Sequence 23, Appli Sequence 20, Appli Sequence 24, Appli Sequence 1601, Appli Sequence 7, Appli Sequence 1179, Appli Sequence 5, Appli Sequence 3, Appli Sequence 104, Appli Sequence 104, Appli Sequence 104, Appli Sequence 35, Appli Sequence 1, Appli
6	98.8	7.7	123	10	US-09-844-864-19 Sequence 1, Appli Sequence 18, Appli Sequence 22, Appli Sequence 21, Appli Sequence 25, Appli Sequence 19, Appli Sequence 23, Appli Sequence 20, Appli Sequence 24, Appli Sequence 1601, Appli Sequence 7, Appli Sequence 1179, Appli Sequence 5, Appli Sequence 3, Appli Sequence 104, Appli Sequence 104, Appli Sequence 104, Appli Sequence 35, Appli Sequence 1, Appli
7	98.8	7.7	123	10	US-09-844-864-23 Sequence 1, Appli Sequence 18, Appli Sequence 22, Appli Sequence 21, Appli Sequence 25, Appli Sequence 19, Appli Sequence 23, Appli Sequence 20, Appli Sequence 24, Appli Sequence 1601, Appli Sequence 7, Appli Sequence 1179, Appli Sequence 5, Appli Sequence 3, Appli Sequence 104, Appli Sequence 104, Appli Sequence 104, Appli Sequence 35, Appli Sequence 1, Appli
8	72	5.6	105	10	US-09-844-864-20 Sequence 1, Appli Sequence 18, Appli Sequence 22, Appli Sequence 21, Appli Sequence 25, Appli Sequence 19, Appli Sequence 23, Appli Sequence 20, Appli Sequence 24, Appli Sequence 1601, Appli Sequence 7, Appli Sequence 1179, Appli Sequence 5, Appli Sequence 3, Appli Sequence 104, Appli Sequence 104, Appli Sequence 104, Appli Sequence 35, Appli Sequence 1, Appli
9	70.4	5.5	105	10	US-09-844-864-24 Sequence 1, Appli Sequence 18, Appli Sequence 22, Appli Sequence 21, Appli Sequence 25, Appli Sequence 19, Appli Sequence 23, Appli Sequence 20, Appli Sequence 24, Appli Sequence 1601, Appli Sequence 7, Appli Sequence 1179, Appli Sequence 5, Appli Sequence 3, Appli Sequence 104, Appli Sequence 104, Appli Sequence 104, Appli Sequence 35, Appli Sequence 1, Appli
10	49.6	3.9	14800	10	US-09-954-456-1601 Sequence 1, Appli Sequence 18, Appli Sequence 22, Appli Sequence 21, Appli Sequence 25, Appli Sequence 19, Appli Sequence 23, Appli Sequence 20, Appli Sequence 24, Appli Sequence 1601, Appli Sequence 7, Appli Sequence 1179, Appli Sequence 5, Appli Sequence 3, Appli Sequence 104, Appli Sequence 104, Appli Sequence 104, Appli Sequence 35, Appli Sequence 1, Appli
11	49	3.8	7177	12	US-10-033-026-7 Sequence 1, Appli Sequence 18, Appli Sequence 22, Appli Sequence 21, Appli Sequence 25, Appli Sequence 19, Appli Sequence 23, Appli Sequence 20, Appli Sequence 24, Appli Sequence 1601, Appli Sequence 7, Appli Sequence 1179, Appli Sequence 5, Appli Sequence 3, Appli Sequence 104, Appli Sequence 104, Appli Sequence 104, Appli Sequence 35, Appli Sequence 1, Appli
12	49	3.8	7364	10	US-09-954-456-1179 Sequence 1, Appli Sequence 18, Appli Sequence 22, Appli Sequence 21, Appli Sequence 25, Appli Sequence 19, Appli Sequence 23, Appli Sequence 20, Appli Sequence 24, Appli Sequence 1601, Appli Sequence 7, Appli Sequence 1179, Appli Sequence 5, Appli Sequence 3, Appli Sequence 104, Appli Sequence 104, Appli Sequence 104, Appli Sequence 35, Appli Sequence 1, Appli
13	49	3.8	7364	12	US-10-033-026-5 Sequence 1, Appli Sequence 18, Appli Sequence 22, Appli Sequence 21, Appli Sequence 25, Appli Sequence 19, Appli Sequence 23, Appli Sequence 20, Appli Sequence 24, Appli Sequence 1601, Appli Sequence 7, Appli Sequence 1179, Appli Sequence 5, Appli Sequence 3, Appli Sequence 104, Appli Sequence 104, Appli Sequence 104, Appli Sequence 35, Appli Sequence 1, Appli
14	49	3.8	7376	12	US-10-033-026-3 Sequence 1, Appli Sequence 18, Appli Sequence 22, Appli Sequence 21, Appli Sequence 25, Appli Sequence 19, Appli Sequence 23, Appli Sequence 20, Appli Sequence 24, Appli Sequence 1601, Appli Sequence 7, Appli Sequence 1179, Appli Sequence 5, Appli Sequence 3, Appli Sequence 104, Appli Sequence 104, Appli Sequence 104, Appli Sequence 35, Appli Sequence 1, Appli
15	46.6	3.6	1006	9	US-09-946-807-104 Sequence 1, Appli Sequence 18, Appli Sequence 22, Appli Sequence 21, Appli Sequence 25, Appli Sequence 19, Appli Sequence 23, Appli Sequence 20, Appli Sequence 24, Appli Sequence 1601, Appli Sequence 7, Appli Sequence 1179, Appli Sequence 5, Appli Sequence 3, Appli Sequence 104, Appli Sequence 104, Appli Sequence 104, Appli Sequence 35, Appli Sequence 1, Appli
16	46.6	3.6	1006	10	US-09-795-668-104 Sequence 1, Appli Sequence 18, Appli Sequence 22, Appli Sequence 21, Appli Sequence 25, Appli Sequence 19, Appli Sequence 23, Appli Sequence 20, Appli Sequence 24, Appli Sequence 1601, Appli Sequence 7, Appli Sequence 1179, Appli Sequence 5, Appli Sequence 3, Appli Sequence 104, Appli Sequence 104, Appli Sequence 104, Appli Sequence 35, Appli Sequence 1, Appli
17	46.6	3.6	1006	10	US-09-795-668-104 Sequence 1, Appli Sequence 18, Appli Sequence 22, Appli Sequence 21, Appli Sequence 25, Appli Sequence 19, Appli Sequence 23, Appli Sequence 20, Appli Sequence 24, Appli Sequence 1601, Appli Sequence 7, Appli Sequence 1179, Appli Sequence 5, Appli Sequence 3, Appli Sequence 104, Appli Sequence 104, Appli Sequence 104, Appli Sequence 35, Appli Sequence 1, Appli
18	44.2	3.5	2481	10	US-09-894-998-35 Sequence 1, Appli Sequence 18, Appli Sequence 22, Appli Sequence 21, Appli Sequence 25, Appli Sequence 19, Appli Sequence 23, Appli Sequence 20, Appli Sequence 24, Appli Sequence 1601, Appli Sequence 7, Appli Sequence 1179, Appli Sequence 5, Appli Sequence 3, Appli Sequence 104, Appli Sequence 104, Appli Sequence 104, Appli Sequence 35, Appli Sequence 1, Appli
19	44.2	3.5	15872	9	US-09-860-846-1 Sequence 1, Appli Sequence 18, Appli Sequence 22, Appli Sequence 21, Appli Sequence 25, Appli Sequence 19, Appli Sequence 23, Appli Sequence 20, Appli Sequence 24, Appli Sequence 1601, Appli Sequence 7, Appli Sequence 1179, Appli Sequence 5, Appli Sequence 3, Appli Sequence 104, Appli Sequence 104, Appli Sequence 104, Appli Sequence 35, Appli Sequence 1, Appli

R. Polley

RESULT 2
US-09-844-864-18
Sequence 18, Application US/09844864
; Patent No. US2002042926A1
GENERAL INFORMATION:
APPLICANT: Matzuk, Martin
APPLICANT: Wu, Xuemei
TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
FILE REFERENCE: P01925US2 / 09807797 / OTA 99-48
CURRENT APPLICATION NUMBER: US/09/844,864
CURRENT FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: 60/106,020
PRIOR FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: PCT/US99/25209
PRIOR FILING DATE: 1999-10-28
NUMBER OF SEQ ID NOS: 25
SOFTWARE: Patentin version 3.0
SEQ ID NO 18
LENGTH: 814
TYPE: DNA
ORGANISM: mus musculus
US-09-844-864-18

Query	Match	Score	DB	Length
QY	601 CAGGCTGCAGGGCAGGCCGGAGCAGCAGCCACCGGAGGGAAAGCGGTCCCCCAGCCGAGCAGGAGGGATGTT	660	60	814
Db	601 CAGGCTGCAGGGCAGGCCGGAGCAGCAGCCACCGGAGGGAAAGCGGTCCCCCAGCCGAGCAGGAGGGATGTT	660	60	814
QY	541 CCGAGAGAGGTGCCGCCAGGAAGGGAGAGGGAGAGGGAGGGAGGGAGGGAGGGAGGGATGTT	600	60	814
Db	541 CCGAGAGAGGTGCCGCCAGGAAGGGAGAGGGAGAGGGAGAGGGAGGGAGGGAGGGAGGGATGTT	600	60	814
QY	421 GCCCCGGTCTCGTCGGTACCTCTGTGGCCTCTCTCTACTGGAGGTGCGAGACCGTA	420	420	814
Db	421 GCCCCGGTCTCGTCGGTACCTCTGTGGCCTCTCTCTACTGGAGGTGCGAGACCGTA	420	420	814
QY	361 TCCGGGGCGGCCACGCTGCAGGGTGCAGGGGAGGGAGGGAGGGATCGGTCA	360	360	814
Db	361 TCCGGGGCGGCCACGCTGCAGGGTGCAGGGGAGGGAGGGATCGGTCA	360	360	814
QY	301 CTCGGGGCGGCCACGCTGCAGGGTGCAGGGGAGGGAGGGATCGGTCA	300	300	814
Db	301 CTCGGGGCGGCCACGCTGCAGGGTGCAGGGGAGGGAGGGATCGGTCA	300	300	814
QY	241 AGCAGCCGGTGCAGGGCTGCGGTGCAGGTGAAACCCGGCGACGCCCTCGGTGCAGTGTCA	300	300	814
Db	241 AGCAGCCGGTGCAGGGCTGCGGTGCAGGTGAAACCCGGCGACGCCCTCGGTGCAGTGTCA	300	300	814
QY	1021 CCCTTAACGGCCCCATCGGAAAGACTTGTGGAGATGCAAGGACAACGCCCTGTCTG	1080	1080	814
Db	1021 CCCTTAACGGCCCCATCGGAAAGACTTGTGGAGATGCAAGGACAACGCCCTGTCTG	1080	1080	814
QY	1081 CGACAGCACCTTCAGCTCAATAACATCATTTAGTAGTGAGAGTGCAGGAAACGTTCTGCTAGA	1140	1140	814
Db	1081 CGACAGCACCTTCAGCTCAATAACATCATTTAGTAGTGAGAGTGCAGGAAACGTTCTGCTAGA	1140	1140	814
QY	1141 TGGGCTTAATGGATGGACAGTGTACTGGATAAACCTGTGAATAAAGGTATTGCAAACA	1200	1200	814
Db	1141 TGGGCTTAATGGATGGACAGTGTACTGGATAAACCTGTGAATAAAGGTATTGCAAACA	1200	1200	814
QY	1201 TCTCTCATGACAGACAGTGTACTGGATAAACCTGTGAATAAAGGTATTGCAAACA	1260	1260	814
Db	1201 TCTCTCATGACAGACAGTGTACTGGATAAACCTGTGAATAAAGGTATTGCAAACA	1260	1260	814
QY	1261 AAAAAAAAAAAAAA 1277	1277	1277	814
Db	1261 AAAAAA•AAAAAAAAA 1277	1277	1277	814

Qy 663 GCGGATGCAGTCTGGATGCCCTGGAGGAGGAGGCCATGTCCCTGGCAGAGATGGCTCAGGA 722
Db 661 GCGGATGCAGTCTGGATGCCCTGGAGGAGGAGGCCATGTCCCTGGCAGAGATGGCTCAGGA 720

Qy 723 CCCGGGTGATTGGATGCCCTCGAGACCAAGCACGGAGCAGGACAA 782
Db 721 CCCGGGTGATTGGATGCCCTCGAGACCAAGCACGGAGCAGGACAA 780

Qy 783 GGAGCGCTTGCGTTCCAGTT 803
Db 781 GCGGAGCGGCTTGCGTTCCAGGT 801

RESULT 3
US-09-844-864-22
; Sequence 22, Application US/09844864
; Patent No. US20020042946A1
; GENERAL INFORMATION:
; APPLICANT: Matzuk, Martin
; APPLICANT: Ren, Yongsheng
; APPLICANT: Wu, Xuemei
; TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
; FILE REFERENCE: P01925US2 / 09807797 / OTA 99-48
; CURRENT APPLICATION NUMBER: US/09/844,864
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: 60/106,020
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: PCT/US99/25209
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 809
; TYPE: DNA
; ORGANISM: mus musculus
; US-09-844-864-22

Query Match 58.9%; Score 752.2; DB 10; Length 809;
Best Local Similarity 97.1%; Pred. No. 2.4e-177;
Matches 778; Conservative 0; Mismatches 18; Indels 5; Gaps 1;

Qy 3 GGGGGGGAGGGGGACGCACCCATGGTCCGGGAGCACGTTCCACCCCTGGCCGA 62
Db 1 GGGGGGGAGGGGGACGCACCCATGGTCCGGGAGCACGTTCCACCCCTGGCCGA 60

Qy 63 TCCTTATCCGAGGCCACCAAAGCGGGATGGCTGGAGGTTCGGAGCCAGGGCTGGCG 122
Db 61 TCCTTATCCGAGGCCACCAAAGCGGGATGGCTGGAGGTTCGGAGCCAGGGCTGGCG 120

Qy 123 ACCCGGCCCCCTCTCCCGGCTACAGACAGCTCATGGCCGGAGTACGTCTGA 182
Db 121 ACCCGGCCCCCTCTCCCGGCTACAGACAGCTCATGGCCGGAGTACGTCTGA 180

Qy 183 CAGCCACCCAGGGCACAGCTCATGGCCCTGCTGGGGATGGCTCCGGACTACGTCTGA 242
Db 181 CAGCCACCCAGGGCACAGCTCATGGCCCTGCTGGGGATGGCTCCGGACTACGTCTGA 240

Qy 243 CAGCCGTGACGCTGGGGTGCAGGTGAACCCGGGAGGCCCTCGTGCAGTCTCACT 302
Db 241 CAGCCGTGACGCTGGGGTGCAGGTGAACCCGGGAGGCCCTCGTGCAGTCTCACT 300

Qy 303 CGGGGGCCACGGCTGGGGTGCAGGTGAACCCGGGAGGCCCTCGTGCAGTCTCACT 362
Db 301 CGGGGGCCACGGCTGGGGTGCAGGTGAACCCGGGAGGCCCTCGTGCAGTCTCACT 360

Qy 363 CTGTCAACCCCGTGGCCACGCCGGGAGATCCCCGGGATCCCTCACTGGAGGTGGGGAGG 422
Db 361 CTGTCAACCCCGTGGCCACGCCGGGAGATCCCCGGGATCCCTCACTGGAGGTGGGGAGG 420

Qy 423 CCCGTTCTCGTCCGTGACCTCTGTGGCCTCTCCACTGGAGGTGGGGAGG 482
Db 421 CCCGTTCTCGTCCGTGACCTCTGTGGCCTCTCCACTGGAGGTGGGGAGG 480

Qy 483 GCAGACACCACGAAGGGAGGGAGGCCATCTCGGGACCCGGGACCCGGAGCC 542
Db 481 GCAGACACCACGAAGGGAGGGAGGCCATCTCGGGACCCGGGACCCGGAGCC 540

Qy 543 GAGAGGGGGCGAGGAAAGCGGTCCCAGCCACCGAGGACCGCATGTCTGGCAGAGATGGCTCAGGA 602
Db 541 GAGAGGGGGCGAGGAAAGCGGTCCCAGCCACCGAGGACCGCATGTCTGGCAGAGATGGCTCAGGA 600

Qy 603 GGCTGCAGGGCAGGGGGGGAGCAGCAGGCCACCACGGAGGACCGCATGTCTGGCAGAGATGGCTCAGGA 722
Db 601 GGCGATGCAGTCTGGGAGCAGCAGGCCACCACGGAGGACCGCATGTCTGGCAGAGATGGCTCAGGA 720

Qy 663 GGGGATGCAGTCTGGGAGGAGGCCATGTCTGGCAGAGATGGCTCAGGA 722
Db 661 GGCGATGCAGTCTGGGAGCAGCAGGCCACCACGGAGGACCGCATGTCTGGCAGAGATGGCTCAGGA 660

Qy 723 CCCGGGTGATTGGATGCCCTCGAGACCAAGCACGGAGCAGGACAA 782
Db 721 CCCGGGTGATTGGATGCCCTCGAGACCAAGCACGGAGCAGGACAA 775

Qy 783 GGAGCGCTTGCGTTCCAGTT 803
Db 776 GGAGCTCCTGGTTCCAGGT 796

RESULT 4
US-09-844-864-21
; Sequence 21, Application US/09844864
; Patent No. US20020042926A1
; GENERAL INFORMATION:
; APPLICANT: Matzuk, Martin
; APPLICANT: Ren, Yongsheng
; APPLICANT: Wu, Xuemei
; TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
; FILE REFERENCE: P01925US2 / 09807797 / OTA 99-48
; CURRENT APPLICATION NUMBER: US/09/844,864
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: 60/106,020
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: PCT/US99/25209
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 21
; LENGTH: 305
; TYPE: DNA
; ORGANISM: mus musculus
; US-09-844-864-21

Query Match 21.4%; Score 273.6; DB 10; Length 305;
Best Local Similarity 95.3%; Pred. No. 8.1e-59;
Matches 282; Conservative 0; Mismatches 14; Indels 0; Gaps 0

Qy 965 TGTCAAAAGTGTAAAGAAGACTAGATGTGGCTGCCAGTCAGATTGCCAACGTGGACCT 1024
Db 10 TTTCAGAGTTGTTAAAGAAGACTAGATGTGGCTGCCAGTCAGATTGCCAACGTGGACCT 1024

Qy 1025 AACCGCCCCATCGGCAAGACTTGTGGAGATGCAAGGACAACGCCCTGCTCTGGAC 1084
Db 70 AACCGCCCCATCGGCAAGACTTGTGGAGATGCAAGGACAATGCTGTGCTCTGGAC 1084

Qy 1085 AGCACCTCAGCTCAAAATACATCATTAGTGGAGTCAGATTGCCAACGTGGACCT 1144
Db 130 AGCACCTCAGCTCAAAATACATCATTAGTGGAGTCAGATTGCCAACGTGGACCT 1144

Qy 1145 GCTAAATGGAATGGACAAGTGAGCTTCTCCCTCTCACCTCTCCCTTCCAATTC 1204
Db 190 GCTAAATGGAATGGACAAGTGAGCTTCTCCCTCTCCCTTCCAATTC 1204

Qy 1205 CATGACAGACAGTGTACTTGGATATAAGCCTGTGAATAAGGTATTGCAACA 1260
Db 250 CATGACAGACAGTGTACTTGGATATAAGCCTGTGAATAAGGTATTGCAACA 1260

RESULT 5
US-09-844-864-25
; Sequence 25, Application US/09844864
; Patent No. US20020042926A1
; GENERAL INFORMATION:
; APPLICANT: Matzuk, Martin
; APPLICANT: Ren, Yongsheng
; APPLICANT: Wu, Xuemei
TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
FILE REFERENCE: P01925US2 / 09807797 / OTA 99-48
CURRENT APPLICATION NUMBER: US/09/844,864
CURRENT FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: 60/106, 020
PRIOR FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: PCT/US99/25209
PRIOR FILING DATE: 1999-10-28
NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 25
LENGTH: 375
; TYPE: DNA
; ORGANISM: mus musculus
; US-09-844-864-25

Query Match 7.7%; Score 98.8; DB 10; Length 123;
Best Local Similarity 98.0%; Pred. No. 1.3e-15;
Matches 100; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 795 TTTCCAGTCTAGAGCAGAACTACATCACTGCAAGGACTGCAAATCCGGTG 854
Db 9 TGTTCAGTCTAGAGCAGAACTACATCACTGCAAGGACTGCAAATCCGGTG 68

Qy 855 GGAGAGCGCTATGTGTGGTGTGCAGGGCACCAAGTAAGGT 896
Db 69 GGAGAGCGCTATGTGTGGTGTGCAGGGCACCAAGTAAGGT 110

RESULT 6
US-09-844-864-19
; Sequence 19, Application US/09844864
; Patent No. US20020042926A1
; GENERAL INFORMATION:
; APPLICANT: Matzuk, Martin
; APPLICANT: Ren, Yongsheng
; APPLICANT: Wu, Xuemei
TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
FILE REFERENCE: P01925US2 / 09807797 / OTA 99-48
CURRENT APPLICATION NUMBER: US/09/844,864
CURRENT FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: 60/106, 020
PRIOR FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: PCT/US99/25209
PRIOR FILING DATE: 1999-10-28
NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
LENGTH: 123
; TYPE: DNA
; ORGANISM: mus musculus

Query Match 20.4%; Score 260.6; DB 10; Length 375;
Best Local Similarity 96.7%; Pred. No. 1.5e-55;
Matches 266; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 986 AGATGTGCCCTGCCAGTCAGATTGCCACGGCCCTAACGCCCATGGCAAGAC 1045
Db 101 AGATGTGCCCTGCCAGTCAGACCTCGCCACGGTACCTTAGACGCCCATCAGCAAGAC 160

Qy 1046 TTGTGTGGAGATCCAAGCACACCTTCAGCTCAAATAC 1105
Db 161 TTGTGTGAGAGATGCCAGGACAACGCCCTGTCCTGCGCACAGCACCGTCAGCTCCAATAAC 220

Qy 1106 ATCATTAGTGGAGACTGAAACGGTTCTGTAGATGGGAATGGGACAAAGTGA 1165
Db 221 ATGATTTAGTGGAGAGTCGAAAACGTTCTGTAGATGGGGCTAATGGAAATGGACAAGTGA 280

Qy 1166 GCTTTCTCCCCTTCACTTCCAATTCTCATGACAGACAGTGTACTTG 1225
Db 281 GCTTTCTCCCCTTCACTTCCAATTCTCATGACAGACAGTGTACTTG 340

Qy 1226 GATATAAAGCCTGTGAATAAAGGTATTGCAAACA 1260
Db 341 GATATAAAGCCTGTGAATAAAGGTATTGCAAACA 375

RESULT 7
US-09-844-864-23
; Sequence 23, Application US/09844864
; Patent No. US20020042926A1
; GENERAL INFORMATION:
; APPLICANT: Matzuk, Martin
; APPLICANT: Ren, Yongsheng
; APPLICANT: Wu, Xuemei
TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
FILE REFERENCE: P01925US2 / 09807797 / OTA 99-48
CURRENT APPLICATION NUMBER: PCT/US99/25209
CURRENT FILING DATE: 2001-04-27
NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23
LENGTH: 123
; TYPE: DNA
; ORGANISM: mus musculus
; US-09-844-864-23

Query Match 7.7%; Score 98.8; DB 10; Length 123;
Best Local Similarity 98.0%; Pred. No. 1.3e-15;
Matches 100; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 795 TTTCCAGTCTAGAGCAGAACTACATCACTGCAAGGACTGCAAATCCGGTG 854
Db 9 TGTTCAGTCTAGAGCAGAACTACATCACTGCAAGGACTGCAAATCCGGTG 68

Qy 855 GGAGAGCGCTATGTGTGGTGTGCAGGGCACCAAGTAAGGT 896
Db 69 GGAGAGCGCTATGTGTGGTGTGCAGGGCACCAAGTAAGGT 110

RESULT 8
US-09-844-864-20
; Sequence 20, Application US/09844864
; Patent No. US20020042926A1
; GENERAL INFORMATION:
; APPLICANT: Matzuk, Martin
; APPLICANT: Ren, Yongsheng
; APPLICANT: Wu, Xuemei
TITLE OF INVENTION: OVARY SPECIFIC GENES AND PROTEINS
FILE REFERENCE: P01925US2 / 09807797 / OTA 99-48
CURRENT APPLICATION NUMBER: US/09/844,864
CURRENT FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: 60/106, 020
PRIOR FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: PCT/US99/25209
PRIOR FILING DATE: 1999-10-28
NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
LENGTH: 105
; TYPE: DNA
; ORGANISM: mus musculus

Query Match 3.8%; Score 49; DB 10; Length 7364;
 Best Local Similarity 48.0%; Pred. No. 0.016;
 Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY	250 GACGCTGCCGTGCAGTGAACCCGGCGCGACGCCCTCGGTGCAGTGTCACTCGGGCGC 309
Db	2753 GAGGCCCGAAGGCCAGAGGCCGGAGGCCGAGGCCGAGGCCGAGGCCGAGGCCGAGGCCGAC 2812
QY	310 CGCACGCTGCAGCCTGCAGGGTGCAGGCCAGCCCCGAGGCCGAGGCCGAGGCCGAGGCC 369
Db	2813 CGCA-GCCACAGCAAGGAGGCCGAGGCCGAGGCCGAGGCCGAGGCCGAGGCCGAC 2869
QY	370 CCCGTTGGCACGCCGGGGAGATCCCCGGATCCCGCATCCCTCACTGGAGGTTCAGACCGTAGCCCGTC 429
Db	2870 GGCCCAGGGGGAGGCCGAGGCCGAGGCCGAGGCCGAGGCCGAGGCCGAGGCC 2929
QY	430 TCGTCCGTGACCTCTGTGGCTCTCTACTGGAGGTTCAGGCTGCAGACCTAGCCCGTC 489
Db	2930 GAGCGGGAGGCCGAGGCCGAGGCCGAGGCCGAGGCCGAGGCCGAGGCC 2989
QY	490 CCCACGAAGGGAGGGAGGCCGAGGCCGAGGCCGAGGCCGAGGCCGAGGCCGAGACA 549
Db	2990 GGGCCAAAGGGGAGGCCGAGGCCGAGGCCGAGGCCGAGGCCGAGGCC 3049
QY	550 GTGGCCGAGGAAGGGTCCCCAGCCGGAAGCGAGGAAGGGGATGTTCAAGCTGCA 609
Db	3050 GCGGAGGGAGGCCGAGGCCGAGGCCGAGGCCGAGGCCGAGGCC 3109
QY	610 GGGCAGGGGGTGGGAGCAGCACCAGGAGGACCGGAAACAGTGTGGCCGAT 668
Db	3110 CACGAGGCTGTGAGAACCGGAGGAGACCACGGAGAAGGGGCCACGGAGAAGGGAGGCTGAGAT 3168

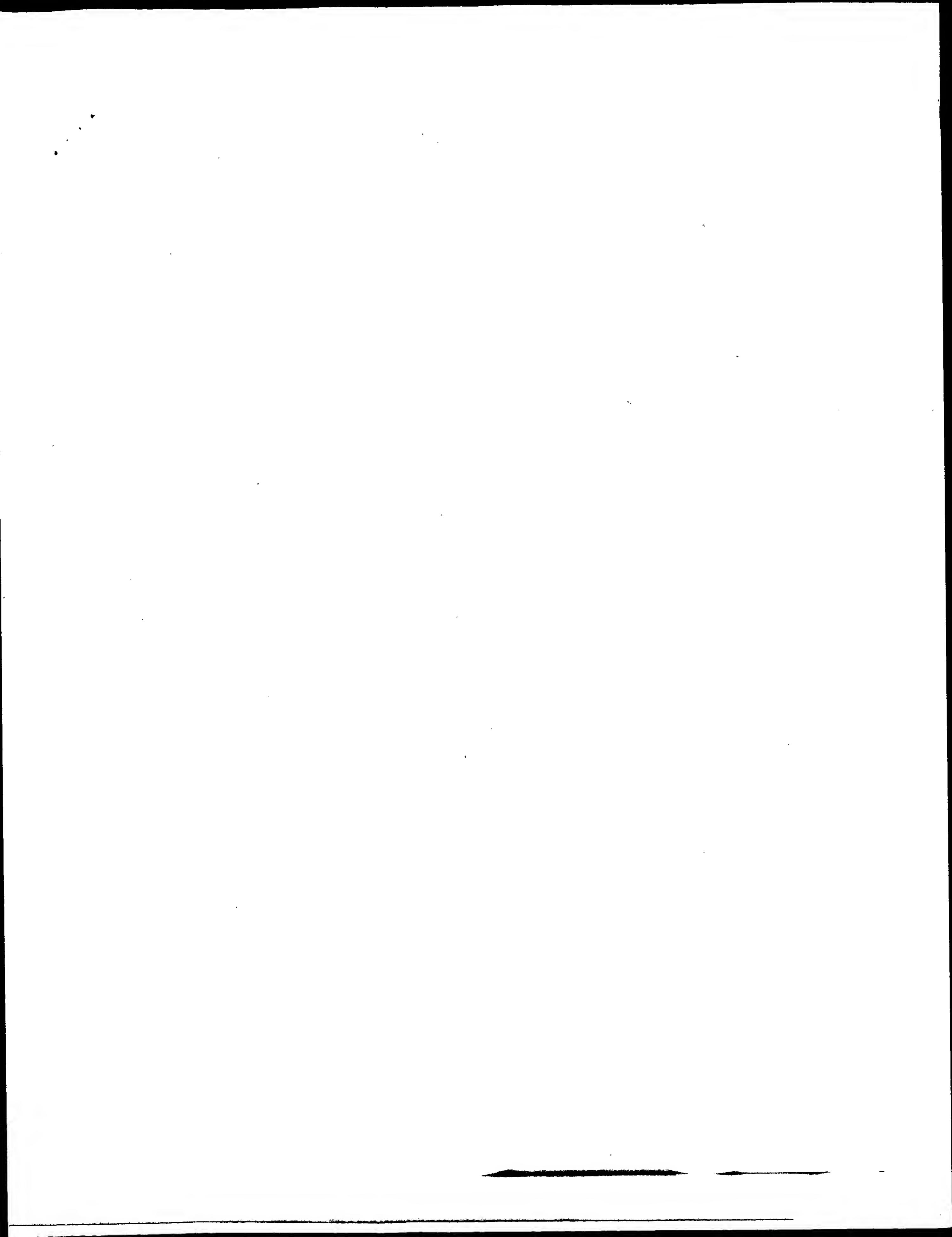
RESULT 12
 US-09-954-456-1179
 ; Sequence 1179, Application US/09954456
 ; Patent No. US20020115057A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Young, Paul
 ; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cane
 ; FILE REFERENCE: 689290-76
 ; CURRENT APPLICATION NUMBER: US/09/954, 456
 ; CURRENT FILING DATE: 2001-09-18
 ; PRIOR APPLICATION NUMBER: US/60/233, 617
 ; PRIOR FILING DATE: 2000-09-18
 ; PRIOR APPLICATION NUMBER: US/60/234, 052
 ; PRIOR FILING DATE: 2000-09-20
 ; PRIOR APPLICATION NUMBER: US/60/234, 923
 ; PRIOR FILING DATE: 2000-09-25
 ; PRIOR APPLICATION NUMBER: US/60/235, 134
 ; PRIOR FILING DATE: 2000-09-25
 ; PRIOR APPLICATION NUMBER: US/60/235, 637
 ; PRIOR FILING DATE: 2000-09-26
 ; PRIOR APPLICATION NUMBER: US/60/235, 638
 ; PRIOR FILING DATE: 2000-09-26
 ; PRIOR APPLICATION NUMBER: US/60/235, 711
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235, 720
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235, 840
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: US/60/235, 863
 ; PRIOR FILING DATE: 2000-09-27
 ; NUMBER OF SEQ ID NOS: 2276
 ; SEQ ID NO 1179
 ; LENGTH: 7364
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 146..7162
 US-10-033-026-5

QY	250 GACGCTGCCGTGCAGTGAACCCGGCGCGACGCCCTCGGTGCAGTGTCACTCGGGCGC 309
Db	2753 GAGGCCCGAAGGCCAGAGGCCGGAGGCCGAGGCCGAGGCCGAGGCCGAGGCCGAC 2812
QY	310 CGCACGCTGCAGCCTGCAGGGTGCAGGCCAGCCCCGAGGCCGAGGCCGAGGCC 369
Db	2813 CGCA-GCCACAGCAAGGAGGCCGAGGCCGAGGCCGAGGCCGAGGCCGAC 2869
QY	370 CCCGTTGGCACGCCGGGGAGATCCCCGGATCCCGCATCCCTCACTGGAGGTTCAGGCTGCAGACCTAGCCCGTC 429
Db	2870 GGCCCAGGGGGAGGCCGAGGCCGAGGCCGAGGCCGAGGCCGAGGCC 2929
QY	430 TCGTCCGTGACCTCTGTGGCTCTCTACTGGAGGTTCAGGCTGCAGACCTAGCCCGTC 489
Db	2930 GAGCGGGAGGCCGAGGCCGAGGCCGAGGCCGAGGCCGAGGCC 2989
QY	490 CCCACGAAGGGAGGGAGGCCGAGGCCGAGGCCGAGGCCGAGGCCGAGGCC 549
Db	2990 GGGCCAAAGGGGAGGCCGAGGCCGAGGCCGAGGCCGAGGCC 3049
QY	550 GTGGCCGAGGAAGGGTCCCCAGCCGGAAGCGAGGAAGGGGATGTTCAAGCTGCA 609
Db	3050 GCGGAGGGAGGCCGAGGCCGAGGCCGAGGCCGAGGCC 3109
QY	610 GGGCAGGGGGTGGGAGCAGCACCAGGAGGACCGGAAACAGTGTGGCCGAT 668
Db	3110 CACGAGGCTGTGAGAACCGGAGGAGACCACGGAGAAGGGGCCACGGAGAAGGGAGGCTGAGAT 3168

RESULT 13
 US-10-033-026-5
 ; Sequence 5, Application US/10033026
 ; Patent No. US20020147309A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lipscombe, Diane
 ; TITLE OF INVENTION: HUMAN N-TYPE CALCIUM CHANNEL ISOFORM AND USES THEREOF
 ; FILE REFERENCE: B1055/7000
 ; CURRENT APPLICATION NUMBER: US/10/033, 026
 ; CURRENT FILING DATE: 2001-12-28
 ; PRIOR APPLICATION NUMBER: US 09/268, 163
 ; PRIOR FILING DATE: 1999-03-12
 ; PRIOR APPLICATION NUMBER: US 60/077, 901
 ; PRIOR FILING DATE: 1998-03-13
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO 5
 ; LENGTH: 7364
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 146..7162
 US-10-033-026-5

Query Match 3.8%; Score 49; DB 12; Length 7364;
 Best Local Similarity 48.0%; Pred. No. 0.016;
 Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY	250 GACGCTGCCGTGCAGTGAACCCGGCGCGACGCCCTCGGTGCAGTGTCACTCGGGCGC 309
Db	2753 GAGGCCCGAAGGCCAGAGGCCGGAGGCCGAGGCCGAGGCCGAGGCC 2812
QY	310 CGCACGCTGCAGCCTGCAGGGTGCAGGCCAGCCCCGAGGCCGAGGCCGAGGCC 369
Db	2813 CGCA-GCCACAGCAAGGAGGCCGAGGCCGAGGCCGAGGCCGAGGCC 2869
QY	370 CCCGTTGGCACGCCGGGGAGATCCCCGGATCCCGCATCCCTCACTGGAGGTTCAGGCTGCAGACCTAGCCCGTC 429
Db	2870 GGCCCAGGGGGAGGCCGAGGCCGAGGCCGAGGCCGAGGCC 2929
QY	430 TCGTCCGTGACCTCTGTGGCTCTCTACTGGAGGTTCAGGCTGCAGACCTAGCCCGTC 489
Db	2930 GAGCGGGAGGCCGAGGCCGAGGCCGAGGCCGAGGCC 2989
QY	490 CCCACGAAGGGAGGGAGGCCGAGGCCGAGGCCGAGGCCGAGGCC 549
Db	2990 GGGCCAAAGGGGAGGCCGAGGCCGAGGCCGAGGCC 3049
QY	550 GTGGCCGAGGAAGGGTCCCCAGCCGGAAGCGAGGAAGGGGATGTTCAAGCTGCA 609
Db	3050 GCGGAGGGAGGCCGAGGCCGAGGCCGAGGCC 3109
QY	610 GGGCAGGGGGTGGGAGCAGCACCAGGAGGACCGGAAACAGTGTGGCCGAT 668
Db	3110 CACGAGGCTGTGAGAACCGGAGGAGACCACGGAGAAGGGGCCACGGAGAAGGGAGGCTGAGAT 3168



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 18, 2003, 09:43:41 ; Search time 296 seconds
(without alignments)
9715.547 Million cell updates/sec

10	49	3.8	7175	16	AAQ84658	Human
11	49	3.8	7175	19	AAV42686	DNA seq.
12	49	3.8	7175	21	AA71704	Human
13	49	3.8	7266	19	AAV29059	Human
14	49	3.8	7362	14	AAQ37817	Sequen.
15	49	3.8	7362	16	AAQ84657	Human
16	49	3.8	7362	19	AAV42685	DNA entry
17	49	3.8	7362	21	AA71703	Human
18	49	3.8	7364	24	ABL65869	Lung cancer
19	49	3.8	7376	20	AAX88001	N-type
20	49	3.8	109519	22	AAS08693	Microm.
21	48.2	3.8	524	22	AAK92108	Human
22	48.2	3.8	524	22	AAK93547	Human
23	48.2	3.8	1769	22	AAK94827	Human
24	47.6	3.7	5944	24	ABL32253	Human
25	47.6	3.7	5944	24	AAQ62837	GCF gene
26	46.6	3.6	745	15	AAQ58330	GGF gene
27	46.6	3.6	745	17	AAT48077	Human
28	46.6	3.6	745	17	AAT30988	Human
29	46.6	3.6	745	17	AAT32061	Human
30	46.6	3.6	745	17	AAT6729	Human
31	46.6	3.6	745	20	AAZ32060	Chemical
32	46.6	3.6	745	20	AAX81175	Human
33	46.6	3.6	745	22	AAQ90317	Human
34	46.6	3.6	745	22	AAK95305	Human
35	46.6	3.6	1006	22	AAK96798	Human
36	46.6	3.6	1986	20	AAQ58329	Human
37	46.6	3.6	1986	22	AAC90318	Human
38	46.6	3.6	2003	15	AAQ62838	Human
39	46.6	3.6	2003	15	AAT48090	Human
40	46.6	3.6	2003	17	AAT30995	L12260
41	46.6	3.6	2003	17	AAT06739	GGF2HBB
42	46.6	3.6	2003	17	AAZ32062	Glial
43	46.6	3.6	2003	20	AAC90319	Glial
44	46.6	3.6	2003	22	AAQ29269	Human
45	45.8	3.6	6232	13	AAQ29269	Human

neuronal cal
ncoding human
calcium chan
calcium chan
ence encoding
neuronal cal
ncoding human
calcium chan
cancer relate
e calcium cha
monospora DNA
CDNA 5'-end
CDNA clone r
full-length
ally treated
immune syste
ne coding se
glial growth
glial growth
glial growth
glial growth
METH2 relate
neuregulin G
CDNA clone
NRG1 exon si
NRG1 exon si
METH2 relate
CDNA clone.
S5. Homo sa
neuregulin G
growth facto
growth facto
METH2 relate
CDNA clone.
calcium chan

CC oligonucleotides (specifically AAX55180-271) can be used for the antisense treatment of diseases and conditions. Typical diseases and conditions are those associated with impaired respiration and inflammation, including lung diseases, pulmonary vasoconstriction, respiration, respiratory distress syndrome, pain, cystic fibrosis, pulmonary hypertension, pulmonary vasoconstriction, emphysema, chronic obstructive pulmonary disease (COPD), and cancers such as leukemias, lymphomas, carcinomas e.g. colon cancer, breast cancer, lung cancer, pancreatic cancer, hepatocellular carcinoma, kidney cancer, melanoma, hepatic metastases, as well as all types of cancers which may metastasize or have metastasized to the lungs, including breast and prostate cancer.

XX SQ Sequence 114955 BP; 6071 A; 29417 C; 36712 G; 21328 T; 21427 other;

Query Match	Score	DB	Length
Best Local Similarity	4.1%	20	114955;
Matches	244;	Pred. No.	0.054;
Conservative	50;	Mismatches	456;
		Indels	5;
		Gaps	2
Qy 13 GCGCGGGACGCCATGTTCCCGGGAGCACGTTCACCCCTGCCGCATCCTTATCCG	72		
Db 111059 GCGCGVGGAGGCAGCNNHHNNSCCGCGACAAGGCCAACGGCCGNNHHNNNSCCGGCAGCG	111000		
Qy 73 CAGGCCACCAAAGCCGGATGGCTGGAGGTTGGAGCCAGGGCTGCCGACCCCGGCC	132		
Db 110999 GCGCGAGCGCNHHNNNSCGCCGGCAAGAGGGCCGGCAGGCCNNHHNNNSCCC	110940		
Qy 133 CCCTCTTCTCCCC---GGCTACAGACAGCTCATGGCCGGAGTACGTCAGACAGCCA	188		
Db 110939 CGCGCCGGCCCCGGGAGNNHHNNNSTCGCCGGAGCCGCAGCCCNNHHNN	110880		
Qy 189 CCAGGGGCACAGCTCATGGCCTGCTGCGGGATGGTCCCCGGTCAGCAGCCG	248		
Db 110879 NSGGCGGCCAGGGCAGGCCACCCGGCCACNNHHNNNSGCCCGGGCAGGVGGCCGVGG	110820		
Qy 249 TGACGCTGCGTGCAGGTGAACCCGGG-CGGGACGCCCTCGGTGCAGTGTCACTCGGGC	307		
Db 110819 CNNHHNNNSCGGGCCGGCVCCGVWVGCCCCGNNHHNNNSCCGGACVGGGG	110760		
Qy 308 GCCGCACGCTGCAGGCCAGGGTGCCAGGGCAGCCCCGACGCCGATCGGTTCTGTGTC	367		
Db 110759 GACGGCGACGCCCNHHNNNSGGGCCCCGGCCGVCCCCGCCAGACCGCC	110700		
Qy 368 AACCCCGTGGCCACGCCGGGGAGATCCCCGGATCCTGGCAGACCGTAGCCCCGT	427		
Db 110699 GGCCCCGGCCCCNNHHNNNSGGGGGGCCNNHHNNNSGGGGGGNHHNNNSGGC	110640		
Qy 428 TCTCGTCCGGTACCTTCTGTGGCTCTCTCACTGGAGGTGCGGGAGGGCAGGCAGA	487		
Db 110639 GGGCGGCCNNHHNNNSGGGGGGGGGGGGGGNNHHNNNSGGGGGGNHHNNNSGGC	110580		
Qy 488 CACCCACCGAAGGGAGAGGGAGCCATCCTCGGGACCCGGAACCGGAGCCGAGAG	547		
Db 110579 GGGCGGCCGGGGNNHHNNNSGGGGGGGGGGNNHHNNNSGGGGGGGGGCN	110520		
Qy 548 AGGTGGCCGGAGGAAGCGGTCCCCCAGCCCGGAAGCGAGGAGGGCGATGTTAGGCTG	607		
Db 110519 NHNNNSGGGGGGGGGGNNHHNNNSGGGGGGGGGGNNHHNNNSGGGGGGGGGCN	110460		
Qy 608 CAGGGCAGGCCGGTGGAGCAGCCACCACGGAGGACCGGAACACTGTGGGGCGA	667		
Db 110459 CGGGCCCNHHNNNSGGGGGGGGGGCCNNHHNNNSGGGGGGGGCCNNHHNNNSGGC	110400		
Qy 668 TGCAGTCTGAGGCCCTGGAGGAGGCCATGCTGGCAGAGATGGCTCAGGCCCG	727		
Db 110399 GCGGGGGCCCNHHNNNSGGGGGGGGGGCCNNHHNNNSGGGGGGGGCCCN	110340		
Qy 728 GTGATTGGATGCCCTCGAGACCAGGCCCTCCCG	762		
Db 110339 NHNNNSGGGGGGGGGGCCCNHHNNNSGGCG	110305		

XX	AC	AAV44436;	Db	359 CGCGGGGACGGTGCAGTCAGG	382
XX	DT	09-NOV-1998 (first entry)	RESULT	6	
XX	DE	Mycobacterium tuberculosis antigen xp22 5' DNA.	ID	AAZ19346	
XX	KW	Tuberculosis; infection; diagnosis; antigen; XP22; ss.	ID	AAZ19346 standard; DNA; 400 BP.	
XX	OS	Mycobacterium tuberculosis strain Erdman.	XX	AAZ19346;	
PN	XX	WO9816645-A2.	XX	AAZ19346;	
PD	XX	23-APR-1998.	XX	AAZ19346;	
XX	PF	07-OCT-1997; 97WO-US18214.	XX	AAZ19346;	
PR	XX	13-MAR-1997; 97US-0818111.	XX	AAZ19346;	
PR	XX	11-OCT-1996; 96US-0729622.	XX	AAZ19346;	
PA	XX	(CORTI-) CORIXA CORP.	XX	AAZ19346;	
XX	PI	Campos-Neto A, Dillon DC, Houghton R, Lodes MJ;	XX	AAZ19346;	
XX	PI	Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;	XX	AAZ19346;	
DR	XX	WPI; 1998-251292/22.	XX	AAZ19346;	
PT	XX	New isolated Mycobacterium tuberculosis polypeptides and DNA - used to develop products for the detection of M. tuberculosis infection and diagnosis of tuberculosis .	XX	AAZ19346;	
PT	XX	Claim 4; Page 185-186; 250pp; English.	XX	AAZ19346;	
CC	CC	This is the 5' region of DNA coding for an antigenic portion of Mycobacterium tuberculosis antigen XP22; 3' DNA is provided in AAV44437. XP22 DNA was isolated from a M. tuberculosis strain Erdman genomic DNA expression library using sera from patients having extrapulmonary tuberculosis. It bears no similarity to known sequences. The invention relates to methods and compositions for diagnosing tuberculosis. It provides polypeptides (see AW64291-W64379) comprising an antigenic portion of a soluble M. tuberculosis antigen, or an immunogenic portion of a M. tuberculosis antigen, as well as DNA sequences encoding such polypeptides, recombinant expression vectors and transformed or transfected host cells. Also claimed are methods and diagnostic kits for detecting M. tuberculosis infection in a patient using these polypeptides, antibodies or oligonucleotide probes and primers.	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 19; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; Gaps 0;	XX	AAZ19346;	
CC	CC	Sequence 400 BP; 74 A; 151 C; 162 G; 13 T; 0 other;	XX	AAZ19346;	
CC	CC	Query Match 3.9%; Score 49.6; DB 20; Length 400; Best Local Similarity 49.2%; Pred. No. 0.033; Matches 130; Conservative 0; Mismatches 134; Indels 0; G			

CC that allow controlled entry of calcium ions into cells. This leads CC to depolarisation events required for muscle contraction. The recombinant subunit, when expressed with nucleic acids encoding the CC complete calcium channel, can be used in assays for the detection and CC characterisation of compounds that modulate the channel. The DNA encoding CC the subunits can be alternatively spliced when transcribed, giving more CC slightly different properties. In addition, the reactivity of the alpha 1 CC subunit with IgG molecules from the serum of an individual with Lambert CC Eaton Syndrome (LES) can be used as a diagnostic for the disease.

SQ sequence 7175 BP; 1415 A; 2197 C; 2168 G; 1395 T; 0 other;

Query Match 3.8%; Score 49; DB 19; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.11; Length 7175;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY 250 GACGCTGGGTGCAGGTGAACCGCGGCCGAGCCTCGGTGCAGTGTCACTCGGGCGC 309

Db 2751 GAGGGCCCCGAAAGGGAGAGGGGGAGGGAGCCGGTCCCTGTCAA 2810

QY 310 CGCACGCTGCAGGCTGCAGGGTGCAGCCAGGCCCCGATCGGTTCCTGTCAA 369

Db 2811 CGCA-GCCACAGCAAGGAGGGGGGGCC-CCGGAGGGAGCGCGGGCGA 2867

QY 370 CCCCGTGGCACGGCCGAGGATCCCCGATCCTGGCAGACGGTAGCCCCGTT 429

Db 2868 GGGCCAGGGCCCGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2927

QY 430 TCGTCCGTGACCTCTGTGCGCTCTCTACTGGAGGTGCGGGAGGCAGACA 489

Db 2928 GAGGGGGAGGCCGACGCCACCGGGACAGGATTCGAGGTGGCC 2987

QY 490 CCCACGAAGGGAGAGGGGGGGCATCTCTGGGAGCCGGGAGCCGGAGAG 549

Db 2988 GGGCCCAGGGCGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 3047

QY 550 GTGGCCGGAGGAAAGGGTCCCCAACCGCGGAAGGGGGGAGATGTTCAGGCTGCA 609

Db 3048 GCGGAGAGGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 3107

QY 610 GGGCAGGGGGGGGGGGAGCAGCACCCACCACGGAGGACCGGGGAACAGTGTGGGGGAT 668

Db 3108 CACGAGGCTGTGGAGAAGGAGACCGAACCGGAGAAGGGGGCTGAGAT 3166

RESULT 12

ID AAA71704
ID AAA71704 standard; DNA; 7175 BP.

AC AAA71704;

XX DT 22-DEC-2000 (first entry)

XX DE Human calcium channel alpha1B-2 subunit encoding DNA.

XX KW Human; calcium channel; calcium channel subunit; diagnosis;

XX Lambert Eaton Syndrome; calcium channel subunit alpha1B-2; ds.

OS Homo sapiens.

XX Key FH Location/Qualifiers

FT CDS 144..6857

FT /*tag= a /product= "calcium channel alpha1B-2 subunit"

XX US6096514-A.

XX PD 01-AUG-2000.

XX PF 25-MAY-1995; 95US-0450562.

PR 04-APR-1988; 88US-0176899.

PR 02-FEB-1990; 90US-0482384.

PR 08-NOV-1990; 90US-0603751.

PR 30-NOV-1990; 90US-0620250.

PR 15-AUG-1991; 91US-0745206.

PR 10-APR-1992; 92US-0868354.

PR 13-JUL-1992; 92US-0914231.

PR 11-AUG-1993; 93US-0105536.

PR 05-NOV-1993; 94US-0149097.

PR 07-FEB-1994; 94US-0193078.

PR 04-APR-1994; 94US-0223305.

PR 11-AUG-1994; 94US-0290012.

PR 23-SEP-1994; 94US-0311363.

PR 28-SEP-1994; 94US-0314083.

PR 07-NOV-1994; 94US-0336257.

PR 13-MAR-1995; 95US-0404950.

PA (SIBI-) SIBIA NEUROSCIENCES INC.

PI Ellis SB, Williams ME, McCue AF, Harpold MM;

DR WPI; 2000-548230/50.

DR P-PSDB; AAB10573.

PT Human calcium channel beta subunit polynucleotides, useful for

PT producing recombinant eukaryotic cells and for diagnosing Lambert Eaton

Syndrome -

PS Example II; Column 113-128; 153pp; English.

CC This invention describes a novel isolated DNA molecule (I) comprising a CC sequence encoding a beta3-1 subunit of a human calcium channel. Nucleic acid probes comprising 14-30 contiguous nucleotides of CC beta_3 subunit encoding DNA are useful for isolation and cloning of CC calcium channel subunit-encoding DNA. Recombinant eukaryotic cells that CC express heterologous calcium channel are useful for identifying compounds CC that modulate calcium channel activity and in assays for identifying CC agonists and antagonists of calcium channel activity in humans. Human CC calcium channel subunit or eukaryotic cells expressing the channel are CC useful for diagnosing Lambert Eaton Syndrome (LES) in a human. This CC sequence encodes the human calcium channel alpha1B-2 subunit which is CC described in the method of the invention.

SQ Sequence 7175 BP; 1415 A; 2197 C; 2168 G; 1395 T; 0 other;

Query Match 3.8%; Score 49; DB 21; Length 7175;
Best Local Similarity 48.0%; Pred. No. 0.11; Length 7175;
Matches 201; Conservative 0; Mismatches 215; Indels 3; Gaps 2;

QY 250 GACGCTGGGTGCAGGTGAACCGCGGCCGACGGCTCTGGTGCAGGTTCACTCGGGCGC 309

Db 2751 GAGGGCCCCGAAAGGGAGAGGGGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2810

QY 310 CGCACGCTGCAGGCTGCAGGGTGCAGCCAGGCCCCGATCGGTTCCTGTCAA 369

Db 2811 CGCA-GCCACAGCAAGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2867

QY 370 CCCCGTGGCACGGCCGAGGATCCCCGATCCTGGCAGACGGTAGCCCCGTT 429

Db 2868 GGGCCAGGG 2927

QY 430 TCGTCCGTGACCTCTGTGCGCTCTCTACTGGAGGTGCGGGAGGCAGACA 489

Db 2928 GAGGGGGAGGCCAACGGGGCACCGGGCACCGGGACTCTGGGAGCCGGAGAG 549

Db 2988 GGGCCCAGGGCGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 3047

QY 550 GTGGCCGGAGGAAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 609

Db 3048 GCGGAGAGGGAGGG 3107

QY 610 GGGCAGGG 668

Mon Jan 20 09:05:59 2003

us-09-844-864-1.rng

Page 12

Qy 550 GTGGCCGCGAGGAAGCGGTCCTTCCAGCCGCGAAGCGAGGGCGATGTCAGGCTGCA 609
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3048 GCGGAGAGGGGGAGGCCGGCGGGCAGGGCACAAGGGCAGCCTGCT 3107
Qy 610 GGGCAGGGGGGGGGAGCAGCAGCCACCGGAGGACCGGAACAGTGTGGGGCGAT 668
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3108 CACGAGGCTGTGAGAACCGGAGAACGGCCACGGAGAAGGGAGGTGAGAT 3166

Search completed: January 19, 2003, 10:25:45
Job time : 508 secs